



**17 February 2020**

**[113-20]**

## **Supporting document 1**

Pregnancy warning labels on packaged alcohol

A review of recent literature

Proposal P1050

# Contents

<b>Executive summary</b>	<b>3</b>
<b>1. Introduction</b>	<b>6</b>
<b>2. Approach to this summary</b>	<b>9</b>
<b>3. Overview of the literature</b>	<b>11</b>
<b>3.1 Study locations</b>	<b>11</b>
<b>3.2 Warnings used in studies</b>	<b>11</b>
<b>3.3 Study methods</b>	<b>12</b>
<b>4. Attention &amp; Recall</b>	<b>14</b>
<b>4.1 Level of attention and recall</b>	<b>14</b>
<b>4.2 Design factors that influence attention</b>	<b>18</b>
4.2.1 Size	18
4.2.2 Location/Placement	20
4.2.3 Colour/Contrast	21
4.2.4 Signal word	22
4.2.5 Pictorials	24
4.2.6 Message length	25
4.2.7 Physical interactivity	25
<b>5. Comprehension</b>	<b>26</b>
<b>6. Judgement</b>	<b>32</b>
<b>7. Behaviour</b>	<b>37</b>
<b>8. Conclusions</b>	<b>42</b>
<b>9. References</b>	<b>43</b>
<b>Appendix A: Method</b>	<b>48</b>

## Executive summary

---

FSANZ undertook a literature review to inform the development of risk management options for a pregnancy warning label on packaged alcoholic beverages. The literature review covered the period from November 2008 to July 2019; being the period from the end of the search for the FSANZ commissioned literature review by Wilkinson et al. (2009). The review was not limited to peer-reviewed papers, but also includes reports from the grey literature, primarily government and non-government organisation (NGO) commissioned studies, as these are highly relevant to Australia and New Zealand populations. A total of 47 studies were included in the review.

We adopted the framework of Argo and Main (2004) to summarise relevant information. They identify five dimensions of warning label effectiveness: attention; reading and comprehension, recall, judgement, and behavioural compliance. Attention and recall are discussed together in this review.

### Attention and recall

To be effective, a warning label has to be noticed. It must draw the attention of a consumer. Prompted awareness of existing pregnancy warning labels across the general Australian and New Zealand public ranged from about 26% to 53% and 25% to 29%, respectively. Prompted awareness was generally higher (around 33% to 74%) for specific populations focussed on women (e.g. women with children, women who are pregnant or planning to have a child or have had a child in the previous 18 months). Experience in the countries with mandated warning labels indicates that the level of awareness of warning labels and recall of their content will increase over time.

The reviewed literature shows trends in awareness with some consumer characteristics. The proportion of populations who are aware of pregnancy warning labels decreases as age increases. Those who drink at higher levels or who drink directly from packaged alcoholic containers were more likely to be aware of pregnancy warning labels than those who drink at lower levels or didn't drink directly from the container. There was also some evidence of those with higher levels of formal education being more likely to be aware of pregnancy warning labels than those with lower levels of formal education.

### Design factors influencing attention

Consumers attention to warning labels is influenced by a range of design factors. These design factors can be manipulated to enhance the noticeability of warning labels such that consumers are more likely to notice the warning.

#### Size

For a pregnancy warning label to be effective it first must be noticed and the consumer direct their attention to it. That the size of an element in a label is related to the attention it receives has been long established in consumer and marketing research. The experimental studies using warning labels on alcohol found that increasing the size of warnings led to an increase in the noticeability of the warning. This was also supported by the findings of qualitative studies. There is likely to be a ceiling effect above which increasing the warning size will have only marginal additional benefit. The size and type of font used impacts its readability with larger fonts being more easily read than smaller fonts. Sentences in all capitals can be

harder to read than those in sentence case. A clear and large font is particularly important for the visually impaired.

## Location/Placement

There were few studies that tested the impact of warning location on attention for alcohol products experimentally. Despite this there was evidence from qualitative studies that supports the general contention that location of a pregnancy warning label on the front of alcoholic beverages would receive quicker and/or more attention than those placed elsewhere on the packaging. This was also supported by the tobacco warning research where many studies have highlighted the greater effectiveness of tobacco warnings when placed on the front of tobacco packages compared with the back and side of packages. Borders have been used to draw attention to a warning. Studies highlighted that the context in which the warning is placed can impact attention, hence the border can be used to distinguish and separate the warning from other information that competes for attention.

## Colour/Contrast

Colour has been used in warnings to enhance the attention they receive. Experimental studies identified in the review have primarily tested red and black options. Using red in a warning can increase the speed at which the warning is identified and also increase the reported level of attention the warning receives. The use of the red pictogram was also considered more noticeable in contrast to the black pictogram. Colour operates as a cue that in combination with an appropriate signal word<sup>1</sup> is perceived as implying a greater hazard than the equivalent signal word in black text. Some colour combinations produce contrast that is difficult to read (e.g. yellow on white), and legibility is reduced when the contrast between characters and the background is low. Dark lettering on a white background, or vice versa, rather than similar shades of a similar colour has been recommended to enhance legibility.

## Signal words

There were no studies identified in the review that experimentally tested the influence of signal words on attention. However, a broader research literature has demonstrated that signal words are important in drawing attention to a warning. Signal words can also connote different levels of hazard. In some circumstances the use of authoritative sources can increase the credibility of warnings, but they may also result in a level of reactance<sup>2</sup> in response to the message. A supplementary search of the literature for use of 'pregnancy warning' or 'pregnancy caution' did not locate any studies.

## Pictorials

Pictorials have been used in warnings to both draw attention to the warning and to convey information. Types of pictorial content include representative drawings, such as the standard pictogram, actual photographs, or more abstract symbols (as often used in road signs). Studies generally find that the addition of a pictorial element to a textual warning enhances the level of attention that the warning receives in comparison with a text only warning. Additionally, pictorial elements can bridge literacy and other educational gaps. Some studies

---

<sup>1</sup> Signal word refers to the use of words like 'WARNING' or 'HEALTH WARNING' at the top/start of warning to increase attention to the warning and to convey an indication of the level of harm from the product.

<sup>2</sup> Reactance is a negative state of arousal that can be triggered when individuals feel some perceived or actual loss of freedom.

have explored graphic warnings (realistic photographs) with pictorial warnings to find that graphic warnings may be more effective in altering judgements, however others have found increased resistance to messages as a reaction to graphic warnings.

## Comprehension

Most of the evidence on the comprehension of pregnancy warning labels in Australia and New Zealand has been conducted on behalf of government and NGOs. The research on the standard pictogram suggests it is well understood by participants across target populations of women of childbearing age and young women, as well as the general population. When the pictogram is red and black it is interpreted more like a warning than when other colour combinations are used.

Comprehension of the voluntary warning statement of *It's safest not to drink while pregnant* has been explored in cross sectional surveys showing varying degrees of comprehension across studies. While some studies found very high levels of comprehension, others have identified a small, but significant proportion of key target populations who interpret the text as meaning *you can drink when pregnant but it is safer not to*. A degree of ambiguity was identified in focus groups where the word 'safest' gave rise to the varying interpretations.

Few other text messages have been tested in Australian and New Zealand populations. However, research findings suggest it is important to personalise the message to make it more relevant, and to avoid using definitive language (*will cause*) about causal connections.

## Judgement

Wilkinson et al. (2009) concluded that the impact of warning messages on judgements was equivocal highlighting results that both increased risk perceptions in some populations, and decreasing risk perceptions in others. The studies we identified showed that warnings can influence judgements participants hold about alcohol, and about its risks. In particular, combinations of graphic warnings with text enhance the risk perceptions of products over the risk perceptions from text only warnings and those without warnings at all. Multiple exposure to the same warning across different situations can lead to stronger beliefs in alcohol as a risk factor in some chronic illnesses. The size of warnings also appeared to impact product evaluations such that larger warnings are more likely to reduce positive product evaluations than smaller warnings.

When considering warning message believability, convincingness and relevance, some types of warnings perform better than others. Positively framed warnings were rated more believable than those using fear appeals and those using numerical evidence. Language such as *increases risk* was also considered more believable than language like *can cause*.

## Behaviour

The literature on the impact of warning labels on behaviour was limited. The experimental studies reviewed indicated that warning labels can have an impact on self-reported intentions to reduce alcohol consumption. Studies also identified other behaviours such as seeking further information, visiting a website, and talking to others about the risks of harm from alcohol. There was no strong evidence to suggest that where warning labels have been mandated there has been an impact on levels of consumption. Researchers typically note that the current mandated warnings do not incorporate relevant design factors to enhance their effectiveness.

# 1. Introduction

---

In October 2018, the Australia and New Zealand Ministerial Forum on Food Regulation (the Forum) noted a Decision Regulation Impact Statement (DRIS) with options for progressing pregnancy warning labels on packaged alcoholic beverages. The Forum requested Food Standards Australia New Zealand (FSANZ) consider mandatory pregnancy warning labelling, and that the warning label should include a pictogram and relevant warning statement.

The Forum reiterated government advice in Australia and New Zealand that 'pregnant women do not consume any alcohol'. They further noted that 'pregnancy warning labels on packaged alcoholic beverages can raise awareness and prompt discussions about the risks of consuming alcohol during pregnancy and may also support the establishment of cultural norms in relation to pregnant women not drinking alcohol' (Australia and New Zealand Ministerial Forum on Food Regulation, 2018).

Warnings are a special type of communication that have several purposes. Laughery and Wogalter (2016) identify four purposes for warnings. They *provide information* about a hazard, and its consequences, and provide instructions to manage the hazard. Warnings also seek to *influence behaviour* through the instructions included in the warning. Warnings act as a *reminder*. Individuals may be aware of the hazard and how to manage it, but they need to be aware of that at the appropriate time; a warning can draw attention to the hazard at the right time. A warning can also serve a broader social purpose contributing to a *safer world*.

A pregnancy warning label as envisioned by the Forum acts across these four purposes. A pregnancy warning label may provide information about the hazard of alcohol when consumed while pregnant, as well as the consequences and instructions to reduce the risk that the hazard poses to the unborn baby. A pregnancy warning label on packaged alcohol can provide that information at an appropriate time being at the time of purchase and also at the time of consumption, so the warning label can act as a reminder to women and others of the hazard. The pregnancy warning label can also influence behaviour, such as reducing alcohol intake or abstaining from consuming alcohol while pregnant, thus reducing or eliminating the risk alcohol poses. In addition, it can influence other behaviours, such as seeking further information and discussing options to manage the risk of alcohol with others. Finally, a pregnancy warning label can assist in creating new social norms around consumption of alcohol while pregnant, and in doing so it may serve a broader social role.

This document reviews and summarises the evidence about the effectiveness of warning labels on packaged alcohol as one thread of evidence to inform the design of pregnancy warning labels. This review does not commence from a blank page, rather it acknowledges and builds upon the various strands of works that have proceeded it. Of particular note is the literature review commissioned by FSANZ: *Alcohol warning labels: Evidence of impact on alcohol consumption amongst women of childbearing age* (Wilkinson et al., 2009). Additionally the evaluation of the voluntary labelling initiative commissioned by the Australian Department of Health (Siggins Miller, 2014, Siggins Miller, 2017) and research commissioned by New Zealand's Health Promotion Agency were also key sources (Rout & Hannan, 2016).

Since December 2011, pregnancy warning labels have been included on packaged alcohol on a voluntary basis in both Australia and New Zealand. Not-for-profit organisations established and funded by the alcohol industry, DrinkWise in Australia and Cheers in New Zealand, have developed a series of logos for alcohol manufacturers to use on their products. The logos provide various options that include the statement 'It's safest not to drink

while pregnant' and a pictogram featuring a silhouetted pregnant woman holding a drinking glass enclosed within a circle with a diagonal strikethrough. The organisational website is also included in some options. Figure 1 provides examples of the various voluntary logos.



**Figure 1: Pregnancy warning labels in New Zealand (left) and Australia (right)**

The second evaluation of the voluntary labelling initiative in Australia found that among the products representing about 75% of the alcohol market, approximately 75% of products carry a pregnancy warning (Siggins Miller 2017). Approximately 82% of the pregnancy warning labels were approximately 0.5cm x 0.7cm in size (an area of 35mm<sup>2</sup>), around 12% were smaller and 6% were larger (Siggins Miller 2017). Approximately 65% of the pregnancy warning labels were on the back of alcohol products, with about 21% on the side and just under 10% on the top or bottom of alcohol products (Siggins Miller 2017).

In New Zealand, Tinawi et al (2018) found 80% of products purposely sampled<sup>3</sup> had some form of pregnancy related warning. The pictogram was found on 46% of products, while a text warning was on 34%. The average size of the pictogram was 45mm<sup>2</sup>, while the average height of text in a text warning was 1.6mm (Tinawi et al 2018). These findings are not from a representative sample of available products, though coverage and size findings are consistent with Siggins Miller (2017) who used a more thorough sampling frame.

In this review we use the term *warning label* as a generic term to refer to a printed text and/or graphical statement affixed to a consumer product that provides information about a hazard related to the product and/or information about how to manage that hazard to reduce risks.

The term *standard pictogram* refers to the pictogram with a silhouette of a pregnant woman holding a drinking glass enclosed within a circle with a diagonal strikethrough (see Figure 1).

<sup>3</sup> Tinawi et al (2018) purposely sampled alcohol beverages by selecting the cheapest 10 wines (5 white and five red), 10 NZ beers in the largest package size (20 or 24 bottles or cans), and cheapest 15 4-packs of RTDs. They selected the cheapest/largest as a proxy for sales volume as sales volume data was unavailable, and previous research had shown the heaviest overall drinkers and those who drank heavily in a short period prefer the cheapest alcohol. They supplemented the sample with 1 imported beer from each of 24 countries. Data was collected in 2016.

The standard pictogram may come in various colour combinations (e.g. black with green strikethrough, monochrome, greyscale).

The phrase *voluntary warning statement* refers to the warning statement of '*It's safest not to drink while pregnant*'. This statement was derived from the Australian guidelines to reduce health risks from drinking alcohol (National Health and Medical Research Council, 2009) and is used as part of voluntary warning label initiatives (see Figure 1).

## 2. Approach to this summary

---

The aim of this literature review was to inform the design and content of pregnancy warning labels for packaged alcohol. Appendix A provides a description of the literature review methodology adopted, while this section provides a brief overview.

In this review, FSANZ followed the approach of Wilkinson et al (2009) by adopting Argo and Main's (2004) conceptualisation of warning label effectiveness. Argo and Main (2004) identify five dimensions of effectiveness that broadly correspond to sequential stages of human information processing when consumers are exposed to warnings. The five dimensions are:

- **Attention:** in order for a warning to have any impact on a consumer it must be noticed; without noticing a warning the information of the warning cannot be engaged with, it cannot be understood.
- **Reading and Comprehension:** after noticing a warning, the consumer needs to read and understand the content of the message. If the message is unclear or has multiple interpretations, or the consumer does not have the literacy skills to read and understand the message it is unlikely to be correctly followed.
- **Recall:** in addition to noticing and then reading and understanding a warning message, its encoding in memory will assist the subsequent retrieval and use of that information at an appropriate time.
- **Judgement:** Warnings provide new information, or remind people, about some hazard associated with a product. Individuals may make new judgements about the risks of consuming the product and evaluate the product in light of the new information provided by the warning.
- **Behavioural Compliance:** Warnings seek to encourage safe behaviours and to discourage unsafe behaviours. Warnings may also encourage behaviours that contribute ultimately to a desired behavioural outcome indirectly. For example, a warning may encourage someone to seek further information or talk with someone else about the warning, these actions may subsequently lead to behaviour change.

FSANZ undertook a systematic search for literature that reports on studies of consumer responses to warnings. Empirical studies from the peer-reviewed literature that reported any measures of the five dimensions of effectiveness were in the scope of the review. This included both quantitative and qualitative studies, and reviews that reported empirical findings. Only English language literature was included and the search covered the period November 2008 to July 2019. Our commencement date was the end date of the review undertaken by Wilkinson et al (2009). The following electronic data bases were searched:

- SocINDEX with full text
- EconLit with Full Text
- Food Science Source
- Food Science and Technology Abstracts
- Medline with full text

Hand searching was also undertaken to identify further papers that may not have been captured through the electronic search. Hand searching included review of reference lists of in scope publications.

Nine systematic and narrative literature reviews were identified during the searches of electronic databases and hand searching. These were included in the literature review.

There is a set of studies and reports that are not published in the peer review literature. These include reports commissioned by government agencies (e.g. Rout & Hannan, 2016; Siggins Miller, 2017), and both public health and industry advocacy groups (Hall & Partners, 2018; GALKAL, n.d.; Quantum Market Research, 2019). It is acknowledged that some of this grey literature may lack the methodological rigour and reporting standards of the peer-reviewed literature. However this literature also includes important studies undertaken with Australian and/or New Zealand populations exploring pregnancy warning labels. Accordingly five reports from grey literature have been included in this literature review.

Two supplementary searches were also undertaken. These related to the influence of signal words in warnings on consumers. The first supplementary search identified four papers that were subsequently included in the literature review. These papers were in-scope for topic and content, however they were published outside the timeframe used for initial search. The second supplementary search focused on the use of the terms 'PREGNANCY WARNING' or 'PREGNANCY ADVICE'. No additional studies were identified through this second supplementary search.

Following the removal of duplicates and screening papers on title and abstract, 59 papers underwent full-text review and quality assessment. The relative quality of each paper was assessed as being low, medium or high. Quality was assessed with reference to five key criteria and two additional criteria for grey literature. The criteria were applied across all types of study design. However the criteria were separately operationalised for qualitative and quantitative studies (see Table A1 in Appendix A). Quality assessment covered:

- Study aims, justification, & theory
- Study design & methods
- Population & sampling
- Analysis
- Reporting
- Peer review (for grey literature only)
- Conflict of interest (for grey literature only)

In total 47 papers were included in the review, comprising

- 33 empirical peer-reviewed papers
- 9 systematic and narrative reviews
- 5 reports from the grey literature.

Data about the study and its findings were extracted from the papers and summarised (See Tables A2, A3 and A4 in Appendix A). The five dimensions of effectiveness were used to organise data extraction, along with information about design and sampling, analysis and reporting.

The five dimensions of effectiveness also provided the structure for write-up of the findings from each study. Higher quality studies, and those that used Australian and/or New Zealand populations, were given more weight than lower quality studies and those that used international populations. The majority of studies reviewed were assessed as being of medium quality. Where the write-up refers to studies that were assessed as low or high quality, the quality level was explicitly included in the narrative. Where no indication of quality is indicated the study was assessed as being of medium quality.

## 3. Overview of the literature

### 3.1 Study locations

The empirical studies reported in the peer reviewed literature and reports from the grey literature were undertaken in a number of countries (Table 1). Just under half (18) used Australian and/or New Zealand samples, with the remaining papers reported on studies carried out with European and UK samples (12), and United States and Canadian samples (8).

**Table 1: Study locations of papers reviewed**

Study Location	Papers included	
Australia and/or New Zealand	18	47%*
Europe	6	16%
UK	6	16%
United States of America	5	13%
Canada	3	8%

\* Base: peer reviewed papers and grey literature.

### 3.2 Warnings used in studies

The papers reviewed covered consumer responses to several types of warnings (Table 2). The vast bulk of these were warnings related to alcoholic beverages, with the majority of papers focussing on pregnancy warnings on alcohol or included pregnancy warnings on alcohol as part of their investigations.

**Table 2: Types of warnings in papers reviewed**

Warning type	Papers included	
Pregnancy warnings on alcohol	21	55%*
Cancer warnings on alcohol	9	24%
Other health warnings on alcohol	14	37%
Driving warnings on alcohol	8	21%
Social warnings on alcohol	3	8%
Warnings on other consumer products (not alcohol)	2	5%

\* Base: peer reviewed papers and grey literature; Multiple response categories – percentages do not sum to 100%

### 3.3 Study methods

Seven papers reported on findings derived from qualitative approaches only, 29 papers reported on findings from quantitative approaches only, and two papers used mixed-method approaches combining both quantitative and qualitative methods.

Of the qualitative method papers, focus groups were the predominant method of data collection, with one paper using in-depth interviews. Qualitative studies typically use purposive sampling, that is sampling that seeks information-rich cases to explore. Criteria are typically used to operationalise the sampling (e.g. age ranges, stage of family etc). Importantly, qualitative studies do not seek to be representative of the population, but rather they seek to explore, in this case, in-depth response to alcohol warning labels to gain a detailed understanding. Often qualitative studies are criticised for their small sample sizes which is valid where findings have been generalised beyond their sample. In this review we have used qualitative studies to illustrate the types of responses that individuals may have to warning labels. They are also used where they support or similar findings from studies designed to produce population estimates or more generalizable findings.

Of the quantitative approaches taken, cross-sectional surveys were most commonly used to derive population level estimates, with samples that sought to be representative of target populations (e.g. women of child bearing age or drinkers). Additionally various designs of between- and within-group experiments were used when testing the impact of different design elements or warning message text on dimensions of warning effectiveness. Some papers also used conjoint and discrete choice experimental designs to estimate consumer choice.

A key limitation of the literature is the reliance on data that is self-reported by the research participant. Where participants were asked to recall their past behaviour or experiences with alcohol warnings, the results may be affected by recall bias. Participants may forget or misremember some of their past experiences and behaviours (and so not report these). Social desirability may also be an issue when discussing alcohol consumption while pregnant. As such some respondents may be reluctant to report behaviours perceived as negative, such as drinking while pregnant, while over-reporting their behaviours viewed as positive, such as reducing alcohol consumption. This bias is more likely to impact findings where there is a high level of interaction with others in the research process, for example in interviews or in focus groups. It is less likely in anonymous situations and where there is little interaction with others, for example in survey research where there is no direct interaction between the researcher and the participant. Researchers will seek to reduce the introduction of this bias in the formation of questionnaires and interview/focus group guides.

Some studies used eye-tracking techniques to objectively record where a participant is looking when provided with some stimuli (e.g. Monk et al 2007; Pham et al 2018; Sillero-Rejon et al 2018). In this way attention to labels and label elements can be captured more objectively than through a self-report question.

In some studies, convenience samples are used. For example, many of the experimental studies exploring the impact of design factors on warning label attention reported in this review use convenience samples (e.g. university students). This may be problematic where the study is seeking to generalise their findings to an audience that is not adequately represented by the sample. However, experimental studies will often include a diversity of participants in order to model the influences of key demographic variables on the phenomena of interest. Such studies provide useful information about the influence of the controlled factors (e.g. size or colour or text), but the influences may not hold beyond the modelled group.

The studies on behaviour typically report behavioural intents or historical behaviours. Such self-reported measures may not translate well to real world situations where behaviour may be contingent on other factors. The use of choice modelling and conjoint designs (e.g. Annunziata et al 2019; Jarvis & Pettigrew 2013) seek to approximate real world choice behaviour by incorporating trade-offs among relevant product characteristics into the experiment design.

## 4. Attention & Recall

---

Attracting the attention of women, and the general community, is a necessary initial step in the process of attending to, and acting upon, a pregnancy warning label. The warning label competes with other visual elements on the label, and needs to cut-through in order to attract attention of the consumer. Attention must initially switch to the warning label, to be noticed, and then be maintained upon the warning, in order for information to be extracted and comprehended. A warning label that is not noticed by its intended audience fails at the first step. Attention is the first dimension that Argo and Main (2004) identify in their 5-dimensions of warning label effectiveness.

Recall refers to the ability to remember and recall the details of a warning message from memory. Once a warning message has been encoded, its information may be available as a future source of guidance in subsequent judgements and decisions (Argo & Main, 2004).

While both attention and recall are different dimensions of effectiveness, they have been discussed, together in the literature review as they are often measured similarly in surveys of consumers awareness of warning labels. Recall is typically measured through the accurate description a participant is able to give of a warning label; an unprompted recall. Attention is measured through confirmation of participants' awareness or having previously seen a particular warning from a prompted display of options; a prompted awareness or recall.

This section initially reports on the level of attention and recall that warning labels have achieved in Australia and New Zealand. This generally draws on cross-sectional surveys to provide an estimate for the level of awareness that a particular population or population of interest has. This section then explores various design factors that can be used to enhance the noticeability of warnings, and thus increase the attention and recall they receive.

### 4.1 Level of attention and recall

This section provides an overview of the studies identified that reported the levels of attention or recall that warning labels receive (Table 3). Studies have used both prompted and unprompted approaches. A prompted approach provides a picture or description of the warning label and asks participants if they are aware or recognise the warning (e.g. Quantum Market Research, 2019; Rout & Hannan, 2016). An unprompted approach would seek a description from the participant in response to a more general question about warning labels or alcohol more generally, which, if the response adequately describes the warning, would be a positive indication of recall (e.g. Rout & Hannan, 2016; Siggins Miller, 2017).

**Table 3: Prompted and unprompted awareness of pregnancy warning labels**

Study	Population	Unprompted		Prompted	
		Pict.	Text	Pict.	Text
GALKAL (n.d.)	Women aged 18-40 yrs (AU)	n/a		74%	
GALKAL (n.d.)	All respondents (AU)	n/a		45%	
Rout and Hannan (2016)	Women (18-34 yrs) (NZ)	17% <sup>4</sup>		46%	49%
Rout and Hannan (2016)	Women with children (NZ)	13%		30%	36%
Rout and Hannan (2016)	All respondents (NZ)	7%		25%	29%
Coomber et al (2015)	Australians (AU) 18-45	16.1%		37.9%	
Coomber et al (2017)	Australians (AU) 18-45	n/a		52.7%	
Siggins Miller (2017)	Women who are currently pregnant, planning a pregnancy or have a child under 18 months (AU)	6.8%	11.4%	38.9	32.5%
Siggins Miller (2017)	All respondents (AU)	4.9%	8.6%	30.8%	25.9%
Quantum Market Research (2019)	Women (AU) 18-44	n/a		59%	
Quantum Market Research (2019)	Pregnant, breastfeeding or planning pregnancy	n/a		56%	
Quantum Market Research (2019)	Australians (AU) 18+	n/a		40%	
Dumas et al (2018)	French pregnant and postpartum women	n/a		66.1%	

In their high quality 2016 study of adult New Zealanders, Rout and Hannan (2016), found that 7% were aware of pregnancy warning messages without visual prompting. The level of awareness among young women<sup>5</sup> and women with children<sup>6</sup> were higher, at 17% and 13% respectively. Prompted awareness for three warning messages: 1) the standard pictogram; 2) the voluntary warning statement; and 3) an alternative warning statement of '*Don't drink pregnant*' were 25%, 29% and 19% respectively for the New Zealanders. Young women reported significantly higher levels of awareness than New Zealanders across the three warnings, at 46%, 49% and 29% respectively. Further demographic analysis revealed that awareness of the warning was highest among younger participants, those in higher income

<sup>4</sup> Unprompted awareness for Rout and Hannan (2016) relates to pregnancy warning message, rather than the pregnancy warning specifically (from Table 9).

<sup>5</sup> Young women was defined as women aged 18 – 34 years.

<sup>6</sup> Women with children was defined as all women with children aged 15 years and younger.

households, those with university qualifications, and those with higher or moderate risks of alcohol dependency<sup>7</sup>.

As part of the evaluation of the voluntary labelling initiative to place pregnancy warning labels on alcoholic beverages, Siggins Miller collected data in 2014 and again in 2017 about the level of awareness of the standard pictogram and the voluntary warning statement (Siggins Miller, 2014, Siggins Miller, 2017). They found prompted awareness decreased from 42.2% to 38.9% for the standard pictogram and increased from 26.3% to 32.5% for the voluntary warning statement between 2014 and 2017 for women who were pregnant, planning to become pregnant in the next two years, or had a child under 18 months<sup>8</sup>. While women in this group were aware of warning messages about alcohol, few were aware of the standard pictogram or the voluntary warning statement until they were prompted<sup>9</sup>.

In an online survey of Australian adult drinkers, Coomber et al (2017) reported 52.7% of the 18-45 year olds were aware of any warning label among the DrinkWise suite of warning labels<sup>10</sup>. Note that this will include those individuals not aware of the pregnancy warning labels but aware of other elements of the suite of warnings. The same study found that the approximately two-thirds of 18-45 year olds considered that 'harm to unborn babies' was likely to be a 'definitely true' consequence of consuming alcohol. Those who were aware of any element of the warning suite were 87% more likely to consider that 'harm to unborn babies' was likely to be a 'definitely true' consequence of consuming alcohol than those who were not aware of at least one warning label element (OR=1.87; 95%CI=1.42-2.46; p<0.001).

An earlier online survey by the same research group found approximately 37.9% of Australian adult drinkers were aware of any of the pregnancy warning labels (standard pictogram, voluntary warning statement) (Coomber et al 2015). Older participants (35-45 years) were significantly less likely to be aware of the warning labels than younger participants (18-24 years) (OR=0.31, 95%CI=0.18-0.54, p<0.001). Those who more frequently binge drank (OR=1.40, 95%CI=1.13-1.72, p=0.002), and those who drank directly from the can or bottle (OR=1.75, 95%CI=1.10-2.77, p=0.018) were more likely to be aware of pregnancy warning labels than those who did not.

DrinkWise have commissioned two studies that reported awareness of their pregnancy warning labels (GALKAL, n.d.; Quantum Market Research, 2019). The GALKAL (n.d.) cross-sectional survey was assessed as being of low quality due to the limited detail on methodology and sampling. They found that 45% of adult Australians that had purchased packaged alcohol in the previous 12 months were aware of at least one of the images from their suite of warning labels (see Figure 1). There was an apparent trend of decreasing awareness with increasing age, such that 89% of those aged 18-24 were aware of at least one of the warning labels. This dropped to 80% for those aged 25-30; 68% for those aged 31-35; 53% for those aged 36-40; and 50% for those aged 41-45 years.

<sup>7</sup> Risk of alcohol dependency was defined as women with a score of 3-12 or men with a score of 4-12 on the AUDIT-C scale.

<sup>8</sup> The equivalent data for the total sample were: standard pictogram: 33.3% in 2014 to 30.8% in 2017; voluntary message text: 19.9% in 2014 to 25.9% in 2017.

<sup>9</sup> Unprompted awareness for 'target women': standard pictogram 4.3% in 2014 to 6.8% in 2017; messages on alcohol products: 5.7% in 2014 to 11.4% in 2017. Note that this is not directly comparable to the prompted awareness of the voluntary message statements as it includes a range of recalled text statements.

<sup>10</sup> The DrinkWise suite of warning labels is broader than just pregnancy warnings and also includes two additional warning statements: 'Is your drinking harming yourself or others?' and 'Kids and alcohol don't mix'. The suite includes the logos and format options as shown in Figure 1.

In a similar on-line survey for DrinkWise, Quantum Market Research (2019) reported 40% of Australian adults had seen at least one of the DrinkWise labels from the suite of pregnancy warning labels DrinkWise promote (see Figure 1). They reported that 59% of women aged 18-44 years had seen at least one warning label, and 56% of those who were pregnant, breastfeeding or planning a pregnancy<sup>11</sup> had seen at least one warning label. A similar trend in awareness with age was found with the proportion who reported their awareness decreasing as age increased.

A number of qualitative studies also reported data on participants' awareness and recall of pregnancy warning labels on alcohol. While these studies do not seek to provide population estimates of awareness and recall, as the previous studies have done, qualitative approaches provide understanding around *why* participants' were aware (or not) of warning.

Jones and Gregory (2010) used focus groups with student participants (aged 18-22 years) from a regional Australian university to examine their attitudes and likely impact of four text warnings using examples of international alcohol products as stimuli. Three of the four warnings used as stimuli included warnings to avoid drinking while pregnant. The warnings also covered cancer, drink-driving and operating machinery. Prior to being exposed to the stimuli, participants noted they had previously seen warnings on international products but not on Australian produced products. Note that data collection occurred prior to the launch of DrinkWise voluntary labelling initiative. After being shown the examples of international warnings they were asked if they thought exposure to them would influence their attitudes or behaviours. The majority didn't believe they would be effective as they perceived the warnings were more relevant to those who drank more than them or to some other demographic group. Participants consistently mentioned the formatting of the stimuli stating that the warning label looks like an ingredients list and would not be noticed, contained too much text, or that it was too small to be noticed. Suggested improvements to warning labels were to increase the size and include more colour, use clearer or larger text, and to include pictures, with participants comparing the warnings to graphic cigarette warnings.

Coomber et al (2018) used a qualitative approach with Australian university students aged 18-25 years, to explore their response to the DrinkWise suite of warnings. This included the standard pictogram and the voluntary warning statement. Using stimuli of enlarged versions of the warnings, and also images of alcohol products with warnings, focus groups were used to discuss the warnings. Some participants had reported seeing warnings whilst drinking in the past, while others had not noticed them prior to participating in the focus groups. Upon first viewing the warnings, most participants commented on the small size of the warning relative to the overall product brand labels.

The awareness findings from a study of French postpartum and pregnant women were also included in this review. The French experience is relevant as the standard pictogram in Australia and New Zealand is based on the French pictogram, and is a similar size and is in similar locations on packaged alcohol (Dumas et al 2018; Siggins Miller 2017). The French pictogram has been required on packaged alcohol sold in France since 2007 (Dumas et al. (2018). In a study using a cross-sectional survey design, assessed as high quality, Dumas et al. (2018) estimated awareness of the French pictogram among French postpartum and pregnant women at 66.1%. Drinkers were significantly more likely to be aware of the pictogram than non-drinkers (77.3% versus 54.3%;  $p < 0.001$ ). When drinking behaviour was controlled, older women were less likely to be aware of the pictogram, as were women with a lower level of education and single women.

---

<sup>11</sup> It is unclear from the results slide deck if this group is limited to women or if it includes male partners of pregnant or breastfeeding women or males involved in planning a pregnancy.

These findings regarding attention are generally consistent with those reached in the literature review by Wilkinson et al. (2009). They note that the level of awareness of warning labels in the US had increased over time despite concerns about the noticeability of the warning labels. Younger people were more likely to be aware of labels than were older groups, and those who drank at higher levels or more frequently were more likely to be aware than those who didn't. In focussing on an adolescent and young adult population, the review by Scholes-Balog et al. (2012) also reported increased awareness of pregnancy warning labels following their introduction.

Qualitative studies enabled a more in-depth investigation of the factors underpinning the levels of attention that existing pregnancy warning labels received. These highlighted a range of aspects related to the size, location and lack of noticeability in the manner in which warning labels had been implemented (Coomber et al., 2018; Jones & Gregory, 2010). The following section discusses the design factors that influence attention.

## 4.2 Design factors that influence attention

In their meta-analysis, Argo and Main (2004) found that the presence of 'vividness-enhancing characteristics' in warnings is more likely to attract consumers' attention than when absent. Vividness-enhancing characteristics are design features and physical characteristics that enhance the vividness of the warning and include font size, colour, spacing, level of specificity, symbols, pictures, and warning location. Laughery and Wogalter (2016) refer to these as design factors and have identified: size, location/placement, colour/contrast, signal word, pictorials, message length and physical interactivity. The following sections discuss each design feature in turn drawing on the empirical literature.

### 4.2.1 Size

Laughery and Wogalter (2016) note that 'bigger is generally better', but qualify it is generally the relative size of the warning relative to other displayed information. Size also incorporates font size and the size of pictures or images used in the warning. There is a large research literature within advertising and marketing that have shown that large objects are more likely to be noticed, noticed more quickly and receive more attention than smaller objects (Peschel & Orquin, 2013).

Several studies identified in this literature review explored the impact of size on the dimensions of warning label effectiveness. Some studies used experimental designs so that the influence of size on effectiveness can be measured. There was also a number of qualitative studies where size was a part of the research design. There were a number of studies that explored the impact of warning size on other dimensions of effectiveness, but did not report measures of attention. These studies are discussed in subsequent sections of the literature review (e.g. Al-hamdani & Smith, (2017a); Al-hamdani & Smith, (2017b) report on the impact of warning size on judgement and are discussed in Chapter 6).

Pham et al (2018) combined an online survey and eye tracking approach to measure the impacts of size and colour of warning labels on attention with Australian participants. The stimuli used were the front and back labels of a bottle of wine with the standard pictogram and the 'get the facts' label (see Figure 1). Three experimental conditions were tested: 1) colour: using a red strikethrough in the pictogram and red text in the logo; 2) size: increased size by 50% and 3) colour and size: incorporating both the colour changes and size increases. The control condition used a black and white pictogram and 'get the facts' label in the standard size recommended in DrinkWise guidance. The control and experimental conditions were identical for both the survey and the eye tracking components of the study.

Pham et al. (2018) reported a significant effect in the level of attention as measured by a self-report composite scale, with participants exposed to the colour and size condition reporting the highest level of attention compared to the smaller monochrome control ( $p=0.014$ ; control mean=5.0, colour and size mean=5.4). There was a trend of increasing attention from the control condition to the colour condition to size and finally the colour and size condition.

The eye tracking component of the study provides an accurate measure of what components of the label participants looked at. Pham et al. (2018) recorded the number of fixations, time to first fixation, and duration of fixations. Importantly, the eye-tracking component revealed that not all participants looked at the warning component of the stimuli. Of the control group, 59% of participants looked at the warning in contrast to 81% of participants in the colour and size condition. However there were no significant differences between the control and other treatment groups across time to first fixation, the number of fixations or the fixation duration across the treatment conditions. The sample sizes for the four groups ranged from  $n=11$  to  $n=17$  (total  $n=42$ ) and as such the study would be underpowered to detect anything but very large effects.

Kersbergen and Field (2017) also used an eye-tracking approach to objectively measure the level of attention that participants paid to package areas of existing alcohol packaging on the UK market. Analysis suggested that the warning labels were attended to for longer periods when they were larger in size and less complex<sup>12</sup>, however size was not a manipulated variable as the study used existing UK alcohol packaging. Kersbergen and Field (2017) note that the UK warning labels on average take up less than 5% of the packaging, and attention is roughly proportional to their size.

The font used in warning messages is another articulation of a warning's size. The size and type of font used impacts its readability with larger fonts being more easily read than smaller fonts. Sentences in all capitals can be harder to read than those in sentence case. A clear and large font is particularly important for the visually impaired (Wogalter & Leonard, 1999).

As noted above in qualitative studies by Jones and Gregory (2010) and Coomber et al. (2018) the small size of existing warning labels was considered by participants to reduce the effectiveness of current voluntary warnings.

In their literature review on the effectiveness of alcohol warning labels, Wilkinson et al. (2009) note that alcohol warning labels represent a small proportion of the size of an overall label and they have not been designed for impact. They cite a number of papers from research on tobacco warnings noting that smokers are more likely to recall larger warnings than they are smaller warnings. They further note that larger warnings have impacts on judgements and behavioural intent and these are discussed in later sections of this report.

That the size of an element in a label is related to the attention it receives has been long established in consumer and marketing research (e.g. Peschel & Orquin, 2013; Peschel, Orquin, & Mueller Loose, 2019) This literature search revealed two experimental studies that explored the impact of warning size on the level of attention. Both studies suggest increasing the size of warnings on alcoholic beverages will lead to an increase in the noticeability of the warning (Pham et al., 2018; Kersbergen & Field, 2017). This was also supported by the findings of qualitative studies.

---

<sup>12</sup> Complexity was measured using measures of file sizes for the images used as stimuli.

## 4.2.2 Location/Placement

Laughery and Wogalter (2016) note that, as a general principle, warnings located close to the hazard both physically, and in time, are more likely to be noticed and encoded in memory than warnings that are not located proximally to the hazard. Thus a warning that is located on the alcohol package being referred to when making purchase and/or consumption decisions is more likely to be noticed than a warning that is not located on the alcohol package being referred to. Warnings in the field of vision when looking at an alcoholic beverage (i.e. front of pack) are more likely to capture attention than those that are not directly in the field of vision (i.e. side or back of package, not on primary package).

Coomber et al. (2018) undertook a study with young adults (18-25 years) who consumed alcohol using focus groups (n=26 across 4 focus groups) recruited from an Australian university. The study was focussed on participants' responses to options for pictorial and graphic warnings, but also collected data on their responses to existing warnings. Participants commented that the proposed warnings were more noticeable than the existing ones, and that the placement of existing warnings on the lower back label reduced the visibility of the warning. Others noted that the warning was placed alongside other back label elements (e.g. barcode, ingredients list) which further obscured the information conveyed.

Dossou et al (2017) used in-depth interviews (n=26) with a youth sample (15-29 years) to explore the location of pregnancy warnings on alcohol products in France. As noted above, the French experience is relevant as the standard pictogram has been mandatory since 2007 on alcoholic beverages sold in France. They used examples of a well-known brand of spirit carrying the pictogram as stimuli in the interviews, and asked about the elements that attracted their attention. The pictogram was 'barely mentioned' in the participant's top-5 lists of elements they noticed. When asked why the pictogram had not drawn their attention the participants noted its location (on the back of the bottle), its size (too small) and its context (surrounded by other elements hindering visibility).

One quantitative study, identified through the supplementary search was particularly relevant to the impact of warning location/placement on attention. The study by Laughery et al (1993) used a series of existing alcoholic beverage containers before and after the introduction of the mandatory warning labels in the US. Participants were required to indicate if a warning was on the product or not. Warnings printed on the front were located by participants significantly more quickly than those printed on the left hand side, then the back and then the right hand side ( $p < 0.01$ ). They also found that warning orientation significantly influenced the time taken to notice a warning, where horizontal warnings were noticed more rapidly than vertically orientated warnings ( $p < 0.001$ ). This study also found that when the labels are filled with non-warning information, locating the warnings took longer.

Laughery et al. (1993) explored the impact of using a border around a warning to enhance its noticeability. They found no significant effect of the presence of the border over its absence in the time taken to identify the presence or not of a warning. Subsequent research by Wogalter and Rashid (1998) found that in the context of warning signs, a thick and colourful border was more likely to attract attention than no border or a thin border. However this study was assessed as being of low quality as the no details apart from the abstract were available.

There were few studies identified in our search that tested the impact of warning location on attention for alcohol products experimentally. Despite this, there was evidence that supports the general contention that location of a pregnancy warning label on the front of alcoholic beverages would receive quicker and/or more attention than those placed elsewhere on the packaging. This is also supported by the tobacco warning research. Wilkinson et al. (2009)

cite numerous studies that highlight the greater effectiveness of tobacco warnings when placed on the front of tobacco packages compared to the back and side of packages.

### 4.2.3 Colour/Contrast

Generally colour and/or other forms of contrast are associated with the greater noticeability of warnings (Laughery & Wogalter, 2016). In addition to attracting attention, some colours also convey additional meaning, thus the colour red has been a common indicator of hazard and assists understanding that the warning is, in fact, actually a warning; while green can be ambiguous and confusing in the context of warnings.

Only one study identified as part of the search used colour as a manipulated variable in their experimental design to measure the effect of colour on attention (Pham et al., 2018). Several qualitative studies identified colour was an attribute of warning design that participants discussed and commented upon in respect of attention to warning labels (e.g. Coomber et al., 2018; Hall & Partners, 2018). Other studies explored the impact of colour on other dimensions of effectiveness (e.g. comprehension in Rout & Hannan, 2016) and are discussed in the relevant sections.

In Pham et al.'s (2018) study (described above) participants gave the highest score for attention to the condition that used the red circle and strikethrough in the standard pictogram combined with an size increase of 50% (mean = 5.4, 95%CI 0.1-0.2) compared to the control condition of a smaller, black/grey standard pictogram (mean = 5.0, 95%CI 6.5-11.9). They reported a significant difference in the mean attention score that participants gave to each condition across three conditions and the control ( $p = 0.014$ ). The attention score for the colour condition (mean = 5.1, 95%CI 10.0-15.8) was not significantly higher than the control, though when combined with an increase in size the score was greater than the increase in size alone (mean = 5.2, 95%CI 0.1-0.2). The second part of Pham et al.'s (2018) study using an eye tracking approach, revealed that 81% of participants looked at the warning in the colour and size condition compared to 59% in control condition. However, there were no significant differences in number of fixations, time to first fixation and the duration of fixation across the three conditions and the control.

A number of non-experimental studies also identified colour as an issue related to attention. The Foundation for Alcohol Research and Education (FARE) commissioned Hall & Partners (2018) to investigate consumer understanding and interpretation of DrinkWise Australia's warning labels. The study used a qualitative approach with focus groups in the city and suburbs of Sydney and in Newcastle. The focus groups were single sex, and covered women who were pregnant or trying to conceive; male partners of women who were pregnant or trying to conceive; and women who had one or more pregnant friends or who had children under 3 years. The focus groups were further stratified across age and education levels. Stimuli used in the focus groups were the standard pictogram and the voluntary warning statement and a series of alternative pictograms and statements<sup>13</sup>. Four of the alternative pictograms used a red circle and diagonal strikethrough, the fifth used a more complex pictogram with several duller colours. Participants highlighted the red colour as being eye catching and makes the pictogram stand out when compared to the standard pictogram. Participants did not draw the same conclusion for the more complex pictogram with several duller colours.

Laughery et al. (1993) used an experimental design to test the influence of design factors on attention on alcohol products. The four conditions used in the experiment were colour (red,

---

<sup>13</sup> These are described and shown in more detail in Table 4.

black), pictogram (presence, absence), warning icon (a triangle with exclamation mark) associated with signal word (presence, absence) and border around warning (presence, absence). The control stimuli was a base label across beer, wine and spirits products that included brand information and logos, company addresses and information, and a bar code. The stimuli used the standard US mandated alcohol warning<sup>14</sup> altered with the design features in each condition. Attention was measured by the time it took participants to accurately identify if the stimuli included the warning. Participants more rapidly identified warnings when printed in red (mean = 2.35 seconds) than when the warning was printed in black (mean = 2.53 s,  $p < 0.02$ ). There was a significant interaction between colour and icon such that the inclusion of the icon reduced the time taken to correctly identify a warning, when the warning was in black (from 2.67 s to 2.39 s,  $p < 0.05$ ), but there was no further improvement when the warning was red (from 2.37 s to 2.34 s,  $p > 0.05$ ).

Some colour combinations produce contrast that is difficult to read (e.g. yellow on white), and legibility is reduced when the contrast between characters and the background is low. Dark lettering on a white background, or vice versa, rather than similar shades of a similar colour has been recommended to enhance legibility (Wogalter & Leonard, 1999).

The literature review by Wilkinson et al. (2009) notes the importance of colour in the context of heuristic cues, that is the use of learned knowledge structures in the form of simple decision rules to make judgements. In this context using a signal word (e.g. Warning) in the colour red serves as a cue to consumers which is perceived as implying a greater hazard than the equivalent signal word in black text (Zuckerman & Chaiken, 1998).

Colour has been used in warnings to enhance the attention that it receives. In both experimental studies included in this review red and black were the colours tested. They found that using red in a warning can increase the speed in which the warning is identified and also increase the reported level of attention the warning receives. The use of the red pictogram was also considered more noticeable in contrast to the black pictogram.

#### 4.2.4 Signal word

Signal words are used in warnings to both attract attention and to generally indicate a level of hazard (Laughery & Wogalter, 2016). The literature search did not identify any studies that experimentally tested the impact of varying signal words on alcohol warning labels on level of attention attained. Several qualitative studies used stimuli that incorporated various signal words and reported on the responses of participants. Supplementary searching identified several papers that tested warning signal words. A second supplementary search was also carried out using the term 'pregnancy warning' and 'pregnancy caution' however no studies were identified that used these as a signal word in warnings.

Thomson et al (2012) explored the reaction to three different signal words in their Victorian focus groups of participants aged 16 years and above: 1) warning, 2) health warning and 3) government health warning. The majority of participants found 'health warning' most acceptable while 'government health warning' was criticised as it was perceived as akin to a 'nanny' state or a 'Big Brother' type message. The signal words were included on warnings that covered a range of possible hazards arising from alcohol.

Jones and Gregory (2010) used international stimuli that used signal words of 'warning' 'government warning' and 'government health warning' all in capital letters. While participants

---

<sup>14</sup> GOVERNMENT WARNING: (1) According to the Surgeon General, women should not drink alcoholic beverages during pregnancy because of the risk of birth defects. (2) Consumption of alcoholic beverages impairs your ability to drive a car or operate machinery, and may cause health problems.

considered the stimuli to be ineffective, suggestions for improvement did not include changes to the signal words beyond generic size and location enhancements.

In their study Pettigrew et al. (2014) included two warning statements that began 'Warning: alcohol ...'; others did not include the initial signal word. One of these statements was the best performing statement in terms of believability, convincingness and personal relevance, while the second performed about 7<sup>th</sup> in the set of 12 statements. Importantly the study was not designed to test the presence/absence of the signal word and the statements differed in other characteristics.

Only one stimuli used the signal word 'warning' in Hall & Partners' (2018) focus group study. This was reported to be effective in attracting attention among their sample of Australians aged 18-45 years.

Wogalter et al (1992) tested the impact of a range of signal words on consumer products on consumer's perceived level of hazard. They tested the signal words 1) NOTE, 2) CAUTION, 3) WARNING; 4) DANGER; 5) LETHAL. The control condition were no warning text and no signal words, and a second control of no signal word (though the warning message was still present). In general they found that the presence of a signal word raised hazard perceptions compared to when no signal words was present. Level of attention was not significantly different across the tested conditions.

Wogalter et al (1999) asked participants to rate warnings across alcohol, tobacco and dietary supplement examples with signal words from different sources with respect to credibility and likelihood of compliance. The signal words tested were: 1) control – no signal word; 2) WARNING; 3) GOVERNMENT WARNING; 4) US GOVERNMENT WARNING; 5) FEDERAL GOVERNMENT WARNING; 6) US FEDERAL GOVERNMENT WARNING. The study found that the use of WARNING significantly increased the level of perceived credibility and increased likelihood of compliance over the no signal words control condition. Credibility was further enhanced with the addition of a source prefix to the word WARNING; the longer, and most specific, prefix US FEDERAL GOVERNMENT was considered the most credible and also the one that participants were most likely to comply with.

A second study reported in Wogalter et al. (1999) explored the influence on credibility and likelihood of compliance of prefixes to WARNING from different sources. They tested specific regulatory agencies, scientific and professional agencies and general statements. When WARNING was present participants gave significantly higher ratings for credibility and likelihood to comply than when the signal words was not used. Credibility and likelihood to comply increased with the addition of general prefixes (e.g. HEALTH WARNING; IMPORTANT HEALTH WARNING) and increased again when specific source organisation were used (e.g. AMERICAN MEDICAL ASSOCIATION WARNING). Both general and specific source prefixes were significantly different from the control in ratings, however they were not significantly different from each other.

Laughery and Wogalter (2016) cite research that 'danger' is more likely to attract attention than 'warning' or 'caution' or no signal word. They further note that people generally do not differentiate between 'warning' and 'caution' but consider these terms to connote hazard levels less than 'danger'. In their systematic review of warnings on alcohol, Hassan and Shiu (2018b) note that the specific source of a message (e.g. from medical/health bodies/government source) can lead to increased credibility and increased compliance than less specific signals. Wilkinson et al. (2009) noted that when a signal word is in red, this will trigger heuristic processing based on the learned associated of harm and red and the signal word.

There were no studies that had experimentally tested the influence of signal words on attention identified. However there is broader research literature that has demonstrated signals words are important in drawing attention to a warning. Signal words can also connote different levels of hazard. In some circumstances, the use of authoritative sources can increase the credibility of warnings, but may also result in a level of reactance<sup>15</sup> in response to the message. A supplementary search of the literature for uses of 'pregnancy warning' or 'pregnancy caution' did not identify any studies.

#### 4.2.5 Pictorials

Pictorials have been used in warnings to both draw attention to the warning and to convey information. They may comprise representative drawings, such as the standard pictogram, actual photographs, or more abstract symbols (as often used in road signs). Pictorials have been found to enhance attention (Laughery & Wogalter, 2016). The literature search identified one study that explored the impact of graphic warnings on attention to warnings on alcohol. A number of qualitative studies included various pictorials as part of the stimuli participants responded to.

In a study assessed as low quality due to its limited generalisability, small sample size and exploratory nature, Monk et al (2017) used an eye tracking approach to measure attention to different types of alcohol warning labels. They used a head-mounted eye-tracker that recorded eye-movements and dwell time on different areas of interest on stimuli shown on a computer screen. The stimuli were a series of graphic (e.g. medical images) and neutral images (e.g. cartoon physiology) and associated warning text (e.g. alcohol damages brain functioning). Among a small sample of 22 participants, they found that participants spent significantly longer looking at the image component of the warning than the text (mean dwell time 5.24 s versus 2.78 s). However, there was no significant difference in the dwell time between graphic and neutral images.

Using a questionnaire one week before the eye-tracking component and then on its completion, Monk et al. (2017) collected data on participants' expectations from consuming alcohol. They found that those whose positive expectancies increased after exposure to the warnings spent longer looking at the image than those whose positive expectancies decreased or stayed the same. While noting the small scale of the study, the alcohol related content of pictures may in some circumstances be enough to cue positive alcohol expectancies.

In an earlier study using an experimental approach Laughery et al. (1993) explored the impact of including a pictorial in a warning on the time taken to correctly identify that a stimuli label of an alcoholic beverage included a warning or not. They found that warnings with a pictorial resulted in significantly faster responses (mean = 2.36 s) than when the pictorial was not present (mean = 2.53 s;  $p < 0.01$ ). They found the same result when an icon<sup>16</sup> was included next to the warning signal word (mean = 2.36 s versus 2.52 s,  $p < 0.01$ ). Pictorials have been used in warnings to both draw attention to the warning and to convey information. Types of pictorial content include representative drawings, such as the standard pictogram, actual photographs, or more abstract symbols (as often used in road signs). Studies generally find that the addition of a pictorial element to a textual warning enhances the level of attention that the warning receives in comparison with a text only warning. Additionally,

---

<sup>15</sup> Reactance is a negative state of arousal that can be triggered when individuals feel some perceived or actual loss of freedom.

<sup>16</sup> Laughery et al. (1993) consider the icon separate to a pictorial in their experiment, however we have considered both to be examples of pictorials in this review.

pictorial elements can bridge literacy and other educational gaps (e.g. see Vallance et al 2018).

#### **4.2.6 Message length**

Laughery and Wogalter (2016) note that brevity has been a generally accepted criterion for warnings; warnings should be no longer than necessary to communicate the needed information. They note that it can be assumed that a longer message is less likely to be read but find the literature to be mixed in this regard. No studies that explored message length were identified in the literature search for this review.

#### **4.2.7 Physical interactivity**

Requiring some form of physical interaction with a warning label, such as requiring the removal of a warning label before a product could be consumed has been somewhat successful in enhancing the level of attention given to a warning (Laughery and Wogalter 2016). No studies that explored physical interactivity of warnings were identified in the literature search for this review.

## 5. Comprehension

---

Once noticed a warning label needs to be read and its message understood. Reading and comprehension is the second dimension of effectiveness identified by Argo and Main (2004). If the warning is not read, or read and misunderstood or read and not understood, then the information cannot be acted upon. Consumer comprehension of a warning depends on the characteristics of the message itself, (e.g. how simple or complex the message is) as well as characteristics of the consumer (e.g. level of literacy or motivation to read the message) (Argo & Main, 2004).

Much of the research on reading and comprehension relevant to the Australian and New Zealand context comes from the grey literature that has been commissioned by government agencies (e.g. Rout & Hannan, 2016; Siggins Miller, 2017) or by NGOs (e.g. Quantum Market Research, 2019; Rout & Hannan, 2016). This literature varies in its quality and typically provides limited methodological detail. Generally, this literature is not peer reviewed, and may be subject to various real or perceived conflicts of interest. Despite this these studies report on participants' comprehension of warning labels from Australian and New Zealand samples and often use warning labels that currently exist in the Australian and New Zealand markets. This makes this set of literature particularly relevant to the discussion on reading and comprehension.

In the high quality 2016 on-line cross-sectional survey of adult New Zealanders commissioned by New Zealand's Health Promotion Agency (HPA), participants were asked about three warning labels: 1) the standard pictogram; 2) the voluntary warning statement; and 3) an alternative warning statement of 'Don't drink pregnant' (Rout & Hannan, 2016).

A primary message of 'Don't drink if pregnant/possibly pregnant' was conveyed by the standard pictogram and the alternative warning text for 80% and 76% of participants. A lower level of participants, 54% associated this primary meaning with the voluntary warning statement. For young women and women with children, this primary message was conveyed to 90% and 86% respectively for the standard pictogram, to 82% and 74% respectively for the alternative warning statement, and to 61% and 66% respectively for the voluntary warning statement.

A secondary message of 'Alcohol consumption during pregnancy can or will harm an unborn baby/pregnant mother' was conveyed to fewer participants for each warning label tested: 10%, 36% and 31% for the standard pictogram, the voluntary warning statement and the alternative warning statement respectively. Importantly, 8% of all participants, and 14% of young women and 8% of women with children, reported that the voluntary warning statement conveyed the message 'You can drink when pregnant but it is safer not to'. This was not the case for either the standard pictogram nor the alternative warning statement. Twelve percent of participants responded with 'don't know' or 'nothing' for the meaning of the standard pictogram. This response was reported by 1% or less for the two warning statements tested.

Rout and Hannan (2016) also asked participants to assess the degree to which each warning label showed: 1) a link between pregnant women drinking and harm to an unborn child, and 2) that you shouldn't drink any alcohol while pregnant. Across both these assessments, the pictogram performed best, and did so for a greater proportion of participants (total well<sup>17</sup> 67% for pregnancy and alcohol harm link and 84% for you shouldn't drink any alcohol while

---

<sup>17</sup> Sum of those indicating

pregnant). The voluntary warning statement performed better than the alternative warning statement in terms of the link between drinking and harm to an unborn child (Total well: 49% versus 41% respectively), while the alternative warning statement performed better than voluntary warning statement in terms of showing that you shouldn't drink any alcohol while pregnant (total well 62% versus 56%). The same pattern was observed for Māori, Pacific and Asian participants, though at consistently higher levels than for New Zealand European participants.

Rout and Hannan's (2016) study tested the standard pictogram in four colour options: duotone gold (circle and strikethrough in darker tone); duotone grey, monochrome green; and red and black (circle and strikethrough in red). A large majority of the total sample (97%), of young women (97%) and of women with children (98%) considered the red and black version of the pictogram looked most like a warning.

As part of the Australian evaluation of the voluntary labelling initiative to place pregnancy health warnings on alcohol products, Siggins Miller conducted on-line cross-sectional surveys in 2014 and again in 2017 (Siggins Miller, 2014, Siggins Miller, 2017). The studies sought understanding of two warning labels: 1) the standard pictogram; and 2) the DrinkWise warning incorporating their web address under 'get the facts' along with the voluntary warning statement.

Using an unprompted open-ended question that was subsequently coded, 92.5% of participants reported that the standard pictogram conveyed the message 'Don't drink alcohol when pregnant' in 2014. In 2017, this proportion dropped to 80.4%. The authors do not provide any reasons for this decrease, though they note some additional categories for coding were identified in the evaluation in 2017, and that this may contribute to the reduction. A second message of 'alcohol causes harm to mother or unborn child' was conveyed to 2.3% and 2.2% in 2014 and 2017, respectively.

For the DrinkWise warning, 34.9% in 2014 and 51.7% in 2017 reported the warning conveyed a message of 'Don't drink when pregnant', while 30.4% in 2014, and 26.5% in 2017, reported the warning conveyed a message of 'Alcohol causes harm to unborn child or mother'.

DrinkWise commissioned two online surveys that included questions on participants' understanding of the DrinkWise warning messages (GALKAL, n.d.; Quantum Market Research, 2019). These studies have limited detail about their sampling and methodological approach and were reported in brief PowerPoint style slide deck. GALKAL (n.d.) was assessed as low quality and Quantum Market Research (2019) as medium. Four alternate DrinkWise warning messages were included: 1) the standard pictogram; 2) website address and 'get the facts'; 3) website address and voluntary warning statement; and 4) website address, 'get the facts', the voluntary warning statement, and the standard pictogram. A fifth version included both 1) and 2) in the context of a back package wine label including additional labelling information. Seventy-five percent of respondents reported that they understood what the messages and images mean, and this level increased to 82% among those aged 18-40 years (GALKAL, n.d.). This was a subjective assessment of comprehension and while participants reported they understood the warning, they did not report what that understanding actually was.

Using an open-ended question where responses were subsequently coded, Quantum Market Research (2019) reported that 98% of participants understood the message conveyed by the DrinkWise warning label. This comprised 82% who were coded into a 'Don't drink/safest not to drink while pregnant' category; 7% to 'Alcohol is harmful to the baby'; and 5% to 'Be safe/responsible/don't take risks'. The authors do not provide further detail on the breakdown of the don't drink/safest not to drink while pregnant combined category so that it is unclear if

that includes interpretations that include consumption of small amounts of alcohol as being safe or if such interpretations were excluded. Follow-up with DrinkWise indicated such responses were included in the 'Don't drink/safest not to drink while pregnant' category.

The Hall & Partners (2018) study commissioned by FARE (see Section 4.2.3) investigated consumer understanding and interpretation of the standard pictogram. They found the message conveyed was considered by the participants to be clear and straightforward and often articulated as 'do not drink alcohol while pregnant'. Participants noted the circle and diagonal strikethrough across the silhouette of the woman drinking, and interpreted this as a 'stop sign', a universal sign for prohibition. However, the pictogram did not appear to challenge participants' beliefs of the risks associated with alcohol consumption during pregnancy. A smaller minority, of mainly male participants, considered the pictogram conveyed a softer message: 'that it is advisable/recommended not to drink alcohol while pregnant'. Some participants in the focus groups also commented on the size of the bump in the pictogram and that it could be interpreted as representing a particular stage of pregnancy, but that was not how they interpreted the pictogram. The authors did not suggest that this interpretation was personally held by any participants. Hall & Partners (2018) report that the voluntary warning statement was considered '... ambiguous and weak. It was equated to a polite suggestion to consider not drinking, rather than a clear directive or definitive warning' (Hall & Partners, 2018, p. 25). They highlighted the word 'safest' as leading to this ambiguity.

Hall & Partners (2018) also tested a range of alternate pictograms and warning statements.. A brief summary of the comprehension findings from these is at Table 4.

**Table 4: Comprehension of alternate pictograms and warning statements (Source: Hall & Partners, 2018)**

	<p>Pictogram with the strike through was considered too complex as it was a double negative.</p> <p>Without the strikethrough it removed the double negative, but in doing so lost the easily recognised instant 'prohibition/do not' message.</p>
	<p>Conveyed the message of 'do not drink' linked to the red strikethrough.</p> <p>Could be interpreted as a specific stage of pregnancy.</p>

	<p>Some considered the message too broad. For example some interpreted it related to purchases or being in the vicinity of alcohol.</p> <p>Conveyed that the message applied to a range of alcohol types.</p>
	<p>Conveyed 'do not drink' but the torso distracted from the primary message.</p> <p>Conveyed to some participants that it applied to red wine.</p>
<p>During pregnancy, no amount of alcohol is safe</p>	<p>Conveyed that small amounts of alcohol are not safe.</p> <p>Some questioned whether this definitive statement was factually correct.</p>
<p>Do not drink alcohol when pregnant</p>	<p>Conveys a clear instruction that was same as the pictogram.</p> <p>No reason for not drinking conveyed.</p>
<p>Alcohol causes birth defects, do not drink when pregnant.</p> <p>Do not use if pregnant: alcohol causes birth defects.</p>	<p>Definitive language implied to some that alcohol always causes birth defects, rather than using 'can' or 'may'.</p> <p>This was challenged by some at low levels of alcohol consumption</p>
<p>Drinking any alcohol can harm your unborn baby.</p> <p>Even small amounts of alcohol can harm unborn babies.</p>	<p>Mainly understood as conveying that small amounts have the potential to cause harm.</p> <p>Didn't have the same pushback as 'causes birth defects' by using 'can harm'; seen as less definitive.</p> <p>'Harm' was considered vague by some participants.</p>
<p>This product should not be used when pregnant or breastfeeding</p>	<p>Meaning was considered straightforward, and the extension to breastfeeding was understood.</p>

Warning: Do not use if pregnant or breastfeeding.	Meaning considered straightforward. 'Warning' implied it was a potentially serious cause or concern.
---	--

The circle and strikethrough used in several of the pictograms appears to be an important element. This readily conveys a message of prohibition or do not use, and is readily associated with warnings. The more detailed aspects of the various pictograms tested produced a range of interpretations, some which may limit the effectiveness of a warning – such that it only applied to a certain stage of pregnancy or types of alcohol.

Earlier we noted the impact that colour can have on consumers' attention to a warning, and in the case of the pictogram the colour red further enhanced the interpretation of the pictogram as a warning. Hall & Partners (2018) conclude that the standard DrinkWise pictogram was the 'strongest option' among those tested.

Participants' responses to the alternative statements highlighted that more definitive language, (e.g. 'alcohol causes birth defects') may be problematic compared to when there is some qualification in the statements (e.g. alcohol can cause birth defects'). Such definitive language led some participants to question the scientific veracity of the statement.

Hall & Partners (2018) further found that personalising the message (e.g. 'your baby' rather than 'babies in general'; and 'unborn baby' instead of 'fetus') increased relevance and emotional resonance. Based on a set of design principles developed from their research, Hall & Partners (2018) suggest: 'Any amount of alcohol may harm your unborn baby' as an example of a warning statement. They did not test this option in their study. Participants generally agreed a warning label should include both a pictogram and a warning statement to maximise the effectiveness of the label. It was considered the warning statement should reiterate the pictogram's message to not drink alcohol while pregnant and/or explain why alcohol should not be consumed during pregnancy.

The study by Dossou et al. (2017) introduced in Section 4.2.2 reported that participants generally understood the meaning of the pictogram, i.e. not to drink while pregnant but the lack of explicit information was lamented (e.g. how dangerous is alcohol to the fetus?).

The review by Wilkinson et al. (2009) noted that none of the research papers they reviewed examined whether or not participants were able to understand the information in the warning label. They drew on research broader than packaged alcohol and pregnancy warning labels and reported that English language skills were important, and noted that some more complex tobacco health warnings required higher levels of education for understanding.

Most of the relevant information on the comprehension of pregnancy warning labels in Australia and New Zealand has been conducted on behalf of government and NGOs. The research on the standard pictogram suggests it is well understood by participants across target populations of women of childbearing age and young women, as well as the general population. When the pictogram is red and black it is seen more like a warning than with other colour combinations.

Comprehension of the voluntary warning statement of 'It's safest not to drink while pregnant' has been explored in cross-sectional surveys with varying degrees of comprehension. While some studies found very high levels of comprehension, others have identified a significant, but small proportion of key target populations who interpret the statement as meaning 'you can drink when pregnant but it is safer not to'. A degree of ambiguity was also identified in focus groups where the word 'safest' gave rise to varying interpretations.

Few other statements have been tested in Australia and New Zealand populations. However, research findings suggest it is important to personalise the message to make it more relevant, and to avoid using definitive language (e.g. 'will cause') about causal connections.

## 6. Judgement

---

Judgement is the fourth dimension of effectiveness in Argo and Main's (2004) framework. Warning labels can influence consumers by providing additional product information that can influence consumers' perceptions of risk associated with that particular product. Pregnancy warning labels can provide information that can alter or challenge consumers' held beliefs and judgements about the risks of consuming alcohol when pregnant.

Argo and Main's (2004) conceptualisation of judgement was focussed on risk and harm perception. In this literature review we have also included other judgements that may not be directly related to risk or harm, but are still important in the effectiveness of a warning. For example consumers' evaluation of a warning's credibility and believability will influence its effectiveness. If a warning is seen as not credible it will not be taken seriously (Beltramini, 1998; Pettigrew et al., 2014). There were studies that report on judgements about harm and risk, as well as studies that explored other evaluations of warnings. Both topics of judgements are discussed in this section.

Pettigrew et al. (2014) tested a series of differently constructed cancer warnings on a sample of Australian drinkers through an online survey. The warnings tested comprised 11 cancer warnings and one general health warning. They varied by message frame (positive versus negative), cancer reference (general versus specific) and manner in which cancer risk was communicated (increases risk versus can cause cancer). Three outcomes were measured: believability, convincingness and personal relevance through 5-point scales. Across the cancer statements younger respondents and those with a tertiary education found the statements more believable than those older and less educated. Similar results were found for convincingness, though females and beer and wine drinkers found the statements more convincing than males and spirits drinkers. High risk drinkers were more likely to find the statements more personally relevant than lower risk drinkers, though there were no significant differences for convincingness and believability.

Analysing by message characteristic Pettigrew et al. (2014) found warning statements that were characterised by positive frame (e.g. Reduce your drinking to reduce your risk of cancer) were considered more believable than those that used fear appeals (e.g. Alcohol can cause breast cancer) or numerical evidence (e.g. Alcohol causes 1 in 20 cancer deaths). This result was generally repeated for convincingness and personal relevance, with some findings only significant for males. Exploring the manner in which cancer was communicated, Pettigrew et al. (2014) found statements that used terms 'increases risk' were more believable than those that used 'can cause'. The same pattern held for convincingness, however there were no significant differences in levels of personal relevance. General cancer statements (e.g. Warning: alcohol increases your risk of cancer) were found to be more believable, convincing and personally relevant than specific cancer warnings (e.g. Alcohol increases your risk of bowel cancer).

Another study by the same research team (Pettigrew et al., 2016), which was assessed as being of high quality, exposed participants to one of six cancer warning statements across a range of situational contexts (warning on a product, newspaper advertisement, comments from child following health class, bus-stop billboard, and a doctor). This approach sought to simulate how alcohol warnings could be disseminated through a comprehensive, public health campaign. Believability, convincingness and personal relevance were measured after exposure to the series of warning situations.

All statements were considered believable, convincing and personally relevant. The only score lower than the mid-point of the scale across the three measures and six statements

was for a statement relating to breast cancer and personal relevance. This may be explained by the mixed sample and perceived lack of relevance of breast cancer for men. There were some significant differences between men and women, and for education and age, for a number of the individual statements and the three measures, however generally there were few differences in responses to the statements across socio-demographic factors.

In a study assessed as having low quality, Al-hamdani and Smith (2015) reported an experiment of warning label options on participants' judgements across beer, wine and spirits. Four warning conditions were used:

- 1) control: standard alcohol bottle without warning
- 2) text warning: text warning about liver cancer occupying 25% of the front label area on a standard bottle
- 3) text and image warning: text warning as for condition 2 and an image of liver cancer occupying 50% of the front label area (text and image occupying 25% each)
- 4) warning and plain packaging: same as condition 3 but on a bottle that had been stripped of all branding and promotional information

The study incorporated a within subjects design for alcohol type and a between subjects design for warning. The sample was a convenience sample from two medium-sized Canadian universities and a large hospital; the sample was not limited to students. Measures included product-based perception measured by participants' response to the product and a consumer-based perception measured by participants' associations with desirable personal attributes of those they associate with the product (e.g. confidence, young, masculine). These were measured through a 5-point Likert scale. Using a multiple choice question, participants indicated which warning condition they had been exposed to. Only one message was used across all conditions and alcohol types so this question assessed correct or incorrect warning recognition.

The study found that warning type had a significant effect on both product-based and consumer-based perceptions. The text and image warning condition and the text and image on plain packaged alcohol condition produced significantly lower product-based evaluations across beer (C<sup>18</sup> mean = 3.61; T+I mean = 2.96, T+I+PP mean = 2.54), wine (C mean = 3.60; T+I mean = 2.48; T+I+PP mean = 2.48) and spirits (C mean = 3.88; T+I mean = 1.98; T+I+PP mean = 2.08). The text warning only produced a significantly lower product evaluation for the spirit condition (C mean = 3.88; T mean = 3.19). There were no significant impacts on the beer and wine conditions.

Similar to product-based perceptions, participants who were exposed to text and image warnings on a plain packaged gave significantly lower consumer-evaluations than those without warnings across beer (C mean = 3.62; T+I+PP mean = 2.37), wine (C mean = 3.12; T+I+PP mean = 2.57) and spirits (C mean = 3.38; T+I+PP mean = 2.47). Text and image warning was significant across wine (C mean = 3.12; T+I mean = 2.45) and spirits (C mean = 3.38; T+I mean = 2.39), but not for the beer product. Participants exposed to the text only warning did not significantly differ in their consumer-evaluation across any of the alcohol types, beer, wine and spirits. These results were consistent regardless of the age and sex of the participant.

These results indicate that text and image warnings and text and image warnings on plain packaging were generally effective in reducing product-based and consumer-based

---

<sup>18</sup> C = control; T = text warning; T+I = text and image warning, T+I+PP = text and image warning on plain packaging.

judgements of alcohol products. In contrast, the text only warning condition was not as effective in reducing perceptions, with the only significant effect when displayed on spirits.

In this study, the text warning occupied 25% of the front package label while the text and image condition and text and image on plain packaging both occupied 50% of the front package label. While the only difference between text only and text and image condition is the addition of the image, the result also suggests that increasing the size of the warning label from 25% to 50% of the front label package contributes to the significant effect. The study was a small study with a limited sample of 92 participants who were sampled from a university/hospital community, so findings from this study to other population requires caution.

In a subsequent study with a larger sample size ( $n=440$ ), though still a convenience sample, Al-hamdani and Smith (2017b) explored the impact of graphic warning size on branded and plain packaged bottles on product-based and consumer-based evaluations across beer, wine and spirits. The warning used was a text statement about liver cancer and a corresponding image and the size of the warning conditions were 50% (25% text +25% image), 75% (25% text + 50% image), and 90% (40% text + 50% image) of the front bottle area, excluding neck and shoulder. In the plain packaged conditions branding information, logos, artwork and other distinguishing features were removed and replaced with size 12 font of the brand name and descriptor.

Product-based perceptions, consumer-based perceptions and product boringness were measured in the study. There were no main effects of warning size on the three evaluations. Importantly the study involved three conditions of size which were tested – there was no control of no warning in this study design. A significant interaction between warning size and alcohol type for product-based rating was observed. Wines and spirits with warning size of 90% had significantly lower product evaluations than their 50% sized warning counterparts (wine: mean 90% =1.86, mean 50% = 2.24; spirits: mean 90% 1.78, mean 50% 2.11). This did not hold for beer. This study shows only modest gains in increasing front warning size beyond 50%, with significant effect only for wine and spirits in one measure when increased from 50% to 90%. Drawing on this result and their earlier study (Al-hamdani & Smith, 2015), and the broader experience from the tobacco warning literature Al-hamdani and Smith (2017b) suggest a threshold effect of warning sized around 50% of the front label in reducing product-based and consumer-based perceptions.

Wigg and Stafford (2016) used an experimental approach to measure the impact of text and pictorial warnings on participants' fear arousal, perceptions of health risk and intention to reduce and quit alcohol consumption. Participants were exposed to one of three warning label conditions: 1) control – no warning; 2) text-only – with the statement 'Alcohol causes fatal liver cancer'; and 3) pictorial plus text – the text warning as in condition 2 with an image of a diseased liver. The warnings were on placed on realistic images of beer and wine bottles that otherwise included normal branding, marketing and business information. The warnings used in conditions 2 and 3 were the same size and were featured on the front of the wine/beer bottle. The sample was a small, predominantly female sample recruited from a university ( $n=60$ , 72% female).

There was a significant difference in the participants' perceived risks of alcohol across the three warning conditions ( $p=0.003$ ). Those who were exposed to the pictorial plus text warning perceived higher risks with consuming alcohol than those who were exposed to the control warning. The difference between the text only warning and the control were not significant. There was a significant difference in participants' fear arousal across the three warning conditions ( $p<0.001$ ). Those who were exposed to the pictorial plus text condition reported significantly higher levels of fear arousal than those exposed to the text-only condition and the control condition. There was no significant difference in fear arousal

between those in the text-only or control conditions. These results show that the pictorial plus text warning is more effective in increasing perceptions of health risk and heightening fear of alcohol than the text-only condition and when there was no warning.

Jongenelis et al. (2018) found warning statements significantly increased the extent participants believed in alcohol as a risk factor for various chronic diseases (diabetes, cancer, heart disease and mental illness). The study measured beliefs pre- and post-exposure to a warning statement in a simulation, and only included participants who were considered at risk for long term harm from alcohol. Participants were exposed to a single warning statement five times from five different sources (e.g. billboard, on TV, on alcohol) in three simulated on-line locations (doctor office, at home, at a bus stop).

The study found the largest increases in the extent of belief in alcohol as a risk factor for the particular chronic disease that aligned with the warning. However there was a more general effect of increasing the extent of belief in alcohol as a risk factor for the other diseases suggesting a halo type effect occurring. Notably the change in the extent in the belief in alcohol as a risk factor for liver damage between pre- and post-exposure to the warning was not significant. This was attributed to the high baseline level for the belief in alcohol being a risk factor for liver damage.

In a small study of Italian university students, Annunziata et al (2017) identified three clusters of young students based on their responses to the perceived usefulness and degree of attention to warnings. Warnings that were focussed on drink driving and not to take drink while taking medicine were perceived as being more useful than longer term health risks including impacts on fetal health and brain health. The study highlights the need for warning messages to be targeted to particular groups, rather than relying on generic messages.

Krischler and Glock (2015) undertook a small study in Germany and Luxembourg (n=122) (tested warning labels posed as questions (e.g. Do you really want alcohol to help you loosen your inhibitions?), and as statements (e.g. Yes, alcohol helps you loosen your inhibitions) and a control with no warning on outcome expectancies (both positive and negative) of drinking alcohol among young people. The warning included the question/statement and a photograph of the expectancy (e.g. vomiting in a toilet for the inhibition warnings). They found a significant increase in negative outcome expectancies among those who were allocated to the question group. However, there were significant decreases in positive outcome expectancies, and the statement condition did not significantly impact either positive or negative expectancies.

In a multi-country correlational study that took advantage of the existence of different alcohol harm reduction policies across the European Union, Boluarte et al (2011) looked at the impact of policies on risk perceptions of alcohol among youth. Using existing EU survey data on young people and drugs, they found that of eight policies included, only the presence of blood alcohol limits and requiring warnings on alcohol containers/advertisements were predictive of the level of risk perceptions in youth. The study found that requiring warnings on containers or advertisements almost doubled the probability of individuals indicating a high risk perception towards alcohol compared to where no warnings were required.

Wilkinson et al. (2009) concluded that the impact of warning messages on judgements was equivocal highlighting results that both increased risk perceptions in some populations, and others highlighting decreases. The studies identified in this update showed that warnings can influence the judgements participants had about the products. In particular, combinations of graphic warnings with text enhance risk perceptions of products over risk perceptions from text only warnings and those without warnings at all. Multiple exposures to the same warning across different situations can lead to stronger beliefs in alcohol as a risk factor in some chronic illnesses. The size of warnings also appeared to impact product evaluations such

that larger warnings are more likely to reduce product-based evaluations. However there appears to be a ceiling effect above which further reductions in product evaluations are not apparent.

When considering warning message believability, readability and relevance, some types of warnings are more effective than others. Positive framed warnings were rated more positively than those that used fear appeals and those that used numerical evidence. Language such as 'increases risk' was also considered more believable than language like 'can cause'.

## 7. Behaviour

The fifth dimension of effectiveness identified by Argo and Main (2004) was behavioural compliance. In the context of pregnancy warning labels, the aim for pregnant women and those seeking to become pregnant is to follow current advice not to drink alcohol while they are pregnant. While abstinence is the advice, reducing alcohol consumption or shifting to lower alcohol drink (while maintaining levels of consumption) are also positive behavioural outcomes on the way to cessation. This section summarises the studies that report behavioural changes that arise from warning labels.

Several types of behaviours were reported in the literature reviewed. These include changes in alcohol consumption (reduction in amount consumed, changes in the type of alcohol consumed, stopping the consumption of alcohol), discussions with others about the risks associated with alcohol consumption, interventions by third parties acting on warning information, and seeking further information on risks.

A number of cross-sectional surveys reported on estimates of self-reported behaviours in response to warnings from the voluntary warning initiative. These report past behaviours. In their high quality survey of New Zealanders, Rout and Hannan (2016) reported that the standard pictogram prompted 70% of young women to consider the risks of drinking alcohol while pregnant, 68% considered that the pictogram encouraged them not to drink while pregnant and 37% to talk with a friend or family member about the risks of drinking alcohol while pregnant.

Forty-one percent of those who recalled one of the DrinkWise warnings reported that they 'did something different' as a result (GALKAL n.d.). The majority of these reported they 'shared information with other' (20%) followed by 'reduced alcohol intake' (16%) and then changed type of alcohol (2%). One percent reported they had 'stopped drinking alcohol' as a result of seeing the DrinkWise messaging and pictogram. This grey literature report was assessed as having low quality.

The survey by Quantum Market Research (2019) for DrinkWise reported that 26% of Australians have taken some action as a consequence of seeing one of the DrinkWise warnings. Ten percent reported talking to others about the information, 6% stopped drinking alcohol, 5% reduced alcohol consumption, and 2% reported they changed the type of alcohol drunk<sup>19</sup>. Among women aged 18-44 years, 28% reported actions comprising: talking to others – 11%, stopped drinking alcohol – 10%, reduced alcohol intake – 4%, and no women (18-44) reported changing types of alcohol<sup>20</sup>. Among women who were pregnant, breastfeeding or planning a pregnancy, 37% reported changing behaviour as a result of seeing one of the DrinkWise warnings. These were: stopped drinking alcohol – 12%, reduced alcohol consumption – 11%, talking with others – 9%, changed type of alcohol – 4%<sup>21</sup>.

Coomber et al. (2015) reported use of the DrinkWise website in response to awareness of the 'Get the facts' logo and web address. While none of the participants recalled the logo unprompted, 25.3% of participants recognised the logo once prompted. Approximately 5.9% of the participants reported visiting the website in response to seeing the logo. Those who

<sup>19</sup> 2% were coded as doing something 'other'.

<sup>20</sup> 2% were coded as doing something 'other'

<sup>21</sup> 1% were coded as doing something 'other'.

were aware of the logo were more than 7-times more likely to visit the website than those who did not (OR 7.25;  $p < 0.001$ ), while females were significantly less likely to visit the website than males (OR 0.23;  $p 0.004$ ). More frequent binge drinkers (OR 1.56;  $p 0.022$ ), those who consumed directly from the can or bottle (OR 3.5  $p 0.032$ ), and those who supported health warning labels (OR 4.27;  $p 0.038$ ) were all more likely to visit the website than those who did not.

A number of papers reported on experimental designs to explore the influence of warning on participants behaviours. In contrast to the cross-sectional designs, these typically report on future behaviour or behavioural intent.

Jarvis and Pettigrew (2013) used a discrete choice experiment to explore the contribution a text warning made to consumption decisions for pre-mixed drinks. The experiment manipulated three attributes (brand, alcohol content, warning text) with four levels in each attribute. The warnings chosen were all text based and related to brain health and drink driving. The participants were presented with the scenario of heading out to a social occasion with friends and asked to select one of four possible pre-mixed drinks. Each participant completed the task 16 times. The attributes and choice task context were identified through an initial qualitative phase as being relevant to the study population of young Australian adults (18-25 years).

Jarvis and Pettigrew (2013) reported that brand gave the highest level of usefulness for participants choices, followed by alcohol content and then the warning, with all attributes being significant in influencing their choice. Two of the warning levels were significant (a negative framed message about brain damage reduced choice, while a drink driving message enhanced choice). The sample was not homogenous in how the attributes impacted their choices and subsequent latent class modelling revealed five classes of participants. The analysis indicated that these different classes respond differently to warning messages. They found that a negative framed message had a negative impact on choice for several of the classes identified, and a positive framed message had a positive impact for one of the classes. This suggests that for some people a positive framed message may increase the probability of choosing that product. The importance of this research is that the choice experiment has been more reflective of purchase situations than other self-report approaches to data collection.

Miller et al (2016) used a survey of Australian adults to explore the response to four cancer warnings across a range of possible behaviours (drink less often, encourage friends to drink less often, discuss risks, discuss risks with family/friends, educate children about risk). About 50% or less of the participants agreed that the warning labels would influence their drinking behaviour, but larger proportions of participants agreed the warnings would prompt them to educate their children about cancer risks associated with alcohol. There were higher levels of agreement with specific behavioural intents for warnings that were related to specific cancers than a more general cancer warning. Miller et al. (2016) found that females, those who read labels, and those who reassess behaviour on reading warnings were more likely to agree they would prompt discussions, drink less and educate children, than those who were male, who didn't read labels, and who wouldn't reassess behaviour after reading warnings. Miller et al. (2016) concluded that although warning labels may raise awareness and prompt discussions, they are more limited in impacting drinking behaviour.

In the high quality study described in Section 6, Pettigrew et al. (2016) explored responses to warning statements across situations (described above). The study measured participants' intention to reduce alcohol consumption; the extent to which participants' thought they actually would reduce alcohol consumption, and participants' intention to consume five or more drinks in a single sitting within the next two weeks. These measures were taken before exposure to the warning and again after exposure to the warning. Participants were

significantly more likely to report a perceived need to reduce alcohol consumption and an intention to do so after observing the warning statement than they were prior to their exposure of the warning statements. This was the case for all warning statements, and they were all equally effective. Participants who found the statement more believable and more personally relevant reported larger pre-to post-exposure change in their intentions.

Wigg and Stafford (2016) found significant effects of warning conditions on the behavioural intentions they measured: significantly increased intention to reduce alcohol consumption ( $p=0.038$ ); and increased intention to quit alcohol consumption ( $p=0.025$ ). Participants exposed to the pictorial warning reported significantly higher scores on the intention to reduce alcohol consumption and intention to quit alcohol consumption scales. The difference between the text-only warning and the control were not significant, nor was the difference between the text-only and the pictorial warning.

Jongenelis et al. (2018) (described above) found that warning statements significantly decreased participants intention to consume alcohol following exposure to a warning stating that alcohol increases your risk of one of three chronic diseases (diabetes, cancer, and mental illness). Greatest change was observed for the warning statement '*Alcohol increases your risk of diabetes*'. There were no significant effects for warnings about heart disease and liver damage. This lack of effect was attributed to the higher baseline knowledge of the relationship between alcohol consumption and these conditions.

Annunziata (2019) used a discrete choice experiment to simulate the selection of wine from three options with varying formats of warning labels. The attributes tested were the message, the position of the warning, the size of the warning and the wine's alcohol content. The task scenario was to select a bottle of wine to take to dinner with friends. Both front and back labels were shown and the brand, varietal and cost were held constant. The study found that Generation Y (born between 1978-2000) wine drinkers preferred a wine without a warning with the majority of the choice being driven by negative influence of a long-term health outcome and positive influence of the absence of the warning label. The decision was effected to a lesser extent but still negatively by a front label warning and then a large warning. The study concluded that small, back label warning for long-term health effects, or no warning labels are preferred by the participants. The study highlights the capacity for warnings to impact behaviour. However, other product and consumer attributes may ultimately influence wine choice.

Tam and Greenfield (2010) examined how warning labels influence the likelihood of intervening to change another's behaviour. Their study used data from the evaluation surveys conducted 3.5 and 4.5 years after the introduction of mandatory alcohol warning labels in the US. They found that recall of the drink-driving message had significant and positive direct effect on intervention behaviours. The suite of interventions behaviours tested were: asking someone not to drive, offering to drive someone home, asking someone to take a taxi, trying to take someone's car keys and asking someone to stay at their place. Usual consumption of three or more drinks on a usual drinking day and greater handling of alcoholic beverage containers also had positive direct effects on intervention behaviours. Recall of the warning message also had a positive mediating effect, with those who handled alcohol beverages also recalling the message and subsequently intervening in another's behaviour.

Tam and Greenfield (2010) found that socio demographic characteristics also influence intervention behaviour. Women were more likely than men to intervene, as were those aged 40 years or under compared to older drinkers, and ethnic minorities were more likely to intervene than white respondents. These findings are important because they provide evidence of a mechanism where by individuals who recall warning messages may intervene in another's behaviour to reduce harm to that person. While arguing for further research Tam

and Greenfield (2010) suggest this may occur through reducing the social inappropriateness of intervening in another's behaviour, as the warning message may serve to legitimise otherwise socially challenging interventions.

Behaviour was also discussed in the qualitative studies identified. The high quality focus group study by Coomber (2017) exposing participants to pictorial and graphic warnings reported that the warnings would not make them stop drinking, stating drinking was an activity they did not want to give up. However, participants reported the warnings would make them 'think twice' about drinking too much or 'slow down' or reduce their consumption. They reported that conversations about the warnings were more likely to be 'banter' and 'light-hearted' than serious conversations. Women were less likely to report this approach, but discussion would still be minimal. In this study only two warnings (one graphic and one pictorial) were focussed on pregnancy and the rest focussed on other health consequences of alcohol consumption (cancer, death, brain damage, injury).

In Dossou et al. (2017) study introduced in Section 4.2.2 on the mandated French pictogram, the majority of participants reported that the warnings were not effective in changing behaviour. Participants reported overexposure to warnings generally and the lack of noticeability (size, location, marketing environment) as reasons why they considered the pictogram ineffective.

Using text warnings available on international products, Jones and Gregory (2010) reported that in their focus groups these warnings were unlikely to influence participants in reducing alcohol consumption. Participants highlighted the warnings were not relevant to their personal situations, or that they contained information they already knew. Though others noted that some women were not aware of the risk of drinking during pregnancy.

In their literature review Wilkinson et al. (2009) draw attention to the following behaviours related to the introduction of mandatory warning labels in the US:

- reported increase in the likelihood of respondents having a conversation about the risks of alcohol (Kaskutas & Greenfield, 1992)
- prompted pregnant women to discuss the topic (Kaskutas, Greenfield, Lee, & Cote, 1998)
- the greater number of warning types that respondents were exposed to the more likely they were to discuss alcohol associated risks (Kaskutas & Graves, 1994).

In their meta-analysis Argo and Main (2004) found that warnings moderately influence behavioural compliance. They found that familiarity moderated behavioural compliance the more familiar an individual was with the product the higher the probability of behavioural compliance<sup>22</sup>.

The literature on the effectiveness of pregnancy warning labels on behaviour is limited. The experimental studies reviewed indicated that warning labels have an impact on self-reported intentions to reduce alcohol consumption, and behaviours seeking further information. However, there were no high quality studies identified that demonstrated a decrease in alcohol consumption. Stockwell notes that both alcohol industry researchers and independent researchers 'agree fairly closely that impacts on drinking behaviour are either

---

<sup>22</sup> They also found that the higher the cost of complying with the warning, in terms of effort and time, the less likely that consumers would comply with the warning. However this is primarily in the context of protective and safety equipment and has limited relevance to alcohol warning labels.

non-existent or minimal'. Wilkinson et al. (2009) highlight the deficiencies in the implementation of most warning labels such that they are unlikely to be noticed.

## 8. Conclusions

---

FSANZ undertook a literature review to inform the design and development of the pregnancy warning label. The review confirmed that multiple design elements (size, location, colour, pictorials, signal words) can be used in varying combinations to enhance the noticeability of warning labels. Thus larger, front of pack, warnings using colour, signal words and pictorial elements are likely to attract more attention than warning labels lacking those elements. While some studies have explored the combinations of several design elements (e.g. size and colour, or size and graphic image), none have done so comprehensively. It is likely that some enhancement in attention level can be achieved through the different combinations of the design factors. For example a smaller front of pack warning may be as noticeable as a larger back of pack warning, or a larger black and white warning may be as noticeable as a smaller red warning. This provides risk management with some degree of flexibility to optimise the level of attention that a warning receives.

The literature review identified research on comprehension of existing warning statements and the standard pictogram on alcohol in Australian and New Zealand populations. There was limited research on new warning statements and pictograms. However, while the research findings are not definitive in terms of what statements would work best in Australia and New Zealand, they do provide some guidance for statements. The literature on judgements and behaviour was limited with respect to FASD. While a number of different types of behavioural response have been reported (e.g. changed alcohol consumption patterns, seeking further information, visiting websites, prompting discussions and conversations), these reports are generally self-reported. It is generally accepted that where alcohol warnings have been introduced, they have had limited impact on consumption behaviour. It was also noted that current mandatory warnings in place in other countries have not been designed with a view to optimise the attention they receive.

## 9. References

- Al-hamdani, M. (2014). The case for stringent alcohol warning labels: Lessons from the tobacco control experience. *Journal of Public Health Policy*, 35(1), 65–74. <https://doi.org/10.1057/jphp.2013.47>
- Al-hamdani, M., & Smith, S. (2015). Alcohol warning label perceptions: Emerging evidence for alcohol policy. *Canadian Journal of Public Health = Revue Canadienne De Sante Publique*, 106(6), e395-400. <https://doi.org/10.17269/CJPH.106.5116>
- Al-hamdani, M., & Smith, S. M. (2017a). Alcohol health-warning labels: Promises and challenges. *Journal of Public Health (Oxford, England)*, 39(1), 3–5. <https://doi.org/10.1093/pubmed/fox010>
- Al-hamdani, M., & Smith, S. M. (2017b). Alcohol Warning Label Perceptions: Do Warning Sizes and Plain Packaging Matter? *Journal of Studies on Alcohol and Drugs*, 78(1), 79–87. <https://doi.org/10.15288/jsad.2017.78.79>
- Annunziata, A., Agnoli, L., Vecchio, R., Charters, S., & Mariani, A. (2019). Health warnings on wine labels: A discrete choice analysis of Italian and French Generation Y consumers. *Wine Economics and Policy*, 8(1), 81–90. <https://doi.org/10.1016/j.wep.2019.03.001>
- Annunziata, A., Pomarici, E., Vecchio, R., & Mariani, A. (2016). Do Consumers Want More Nutritional and Health Information on Wine Labels? Insights from the EU and USA. *Nutrients*, 8(7). <https://doi.org/10.3390/nu8070416>
- Annunziata, A., Vecchio, R., & Mariani, A. (2017). Alcohol Warnings and Moderate Drinking Patterns among Italian University Students: An Exploratory Study. *Nutrients*, 9(6). <https://doi.org/10.3390/nu9060628>
- Argo, J. J., & Main, K. J. (2004). Meta-Analyses of the Effectiveness of Warning Labels. *Journal of Public Policy & Marketing*, 23(2), 193–208. <https://doi.org/10.1509/jppm.23.2.193.51400>
- Australia and New Zealand Ministerial Forum on Food Regulation (2018). Final communiqué 11 October 2018. Retrieved from <https://foodregulation.gov.au/internet/fr/publishing.nsf/Content/forum-communique-2018-October>
- Bazzo, S., Battistella, G., Riscica, P., Moino, G., Marini, F., Geromel, M., & Czerwinsky, L. (2012). Evaluation of the impact of the image used in a communication campaign to raise awareness about the effects of alcohol use during pregnancy. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 47(6), 657–662. <https://doi.org/10.1093/alcalc/ags051>
- Beltramini, R. F. (1998). Perceived believability of warning label information presented in cigarette advertising. *Journal of Advertising*, 17, 26–32.
- Boluarte, T. A., Mossialos, E., & Rudisill, C. (2011). The Impact of Alcohol Policies across Europe on Young Adults' Perceptions of Alcohol Risks. *CESifo Economic Studies*, 57(4), 763–788. <https://doi.org/10.1093/cesifo/ifr025>
- Brennan, E., Wakefield, M. A., Durkin, S. J., Jernigan, D. H., Dixon, H. G., & Pettigrew, S. (2017). Public awareness and misunderstanding about DrinkWise Australia: A cross-sectional survey of Australian adults. *Australian & New Zealand Journal of Public Health*, 41(4), 352–357. <https://doi.org/10.1111/1753-6405.12674>
- Brown, K. G., Stautz, K., Hollands, G. J., Winpenny, E. M., & Marteau, T. M. (2016). The Cognitive and Behavioural Impact of Alcohol Promoting and Alcohol Warning Advertisements: An Experimental Study. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 51(3), 354–362. <https://doi.org/10.1093/alcalc/aqv104>
- Cil, G. (2017). Effects of posted point-of-sale warnings on alcohol consumption during pregnancy and on birth outcomes. *Journal of Health Economics*, 53, 131–155.

- Coomber, K., Hayley, A., Giorgi, C., & Miller, P. G. (2017). A Qualitative Investigation of Australian Young Adult Responses to Pictorial and Graphic Alcohol Product Warnings. *Journal of Drug Issues*, 47(4), 622–637. <https://doi.org/10.1177/0022042617719443>
- Coomber, K., Hayley, A., & Miller, P. G. (2018). Unconvincing and ineffective: Young adult responses to current Australian alcohol product warnings. *Australian Journal of Psychology*. (2), 131. <https://doi.org/10.1111/ajpy.12177>
- Coomber, K., Martino, F., Barbour, I. R., Mayshak, R., & Miller, P. G. (2015). Do consumers 'Get the facts'? A survey of alcohol warning label recognition in Australia. *BMC Public Health*, 15, 816. <https://doi.org/10.1186/s12889-015-2160-0>
- Coomber, K., Mayshak, R., Curtis, A., & Miller, P. G. (2017). Awareness and correlates of short-term and long-term consequences of alcohol use among Australian drinkers. *Australian and New Zealand Journal of Public Health*, 41(3), 237–242. <https://doi.org/10.1111/1753-6405.12634>
- Dossou, G., Gallopel-Morvan, K., & Diouf, J.-F. (2017). The effectiveness of current French health warnings displayed on alcohol advertisements and alcoholic beverages. *European Journal of Public Health*, 27(4), 699–704. <https://doi.org/10.1093/eurpub/ckw263>
- Dumas, A., Toutain, S., Hill, C., & Simmat-Durand, L. (2018). Warning about drinking during pregnancy: Lessons from the French experience.
- FSANZ. (2013). Risk analysis in food regulation. Food Standards Australia New Zealand, Canberra. <https://www.foodstandards.gov.au/publications/riskanalysisfoodregulation/Documents/risk-analysis-food-regulation-full-pdf.pdf>
- GALKAL. (n.d.). *DrinkWise label awareness*.
- Hall & Partners. (2018). *Understanding of consumer information messaging on alcohol products: Focus group testing report*. Retrieved from <http://fare.org.au/understanding-of-consumer-information-messaging-on-alcohol-products/>
- Hassan, L., & Shiu, E. (2018a). A systematic review of the efficacy of alcohol warning labels. *Journal of Social Marketing*, 8(3), 333–352. <https://doi.org/10.1108/JSOCM-03-2017-0020>
- Hassan, L., & Shiu, E. (2018b). Communicating Messages About Drinking. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 53(1), 1–2. <https://doi.org/10.1093/alcac/agx112>
- Jarvis, W., & Pettigrew, S. (2013). The relative influence of alcohol warning statement type on young drinkers' stated choices. *Food Quality and Preference*, 28(1).
- Jones, S. C., & Gregory, P. (2010). Health Warning Labels on Alcohol Products—The Views of Australian University Students. *Contemporary Drug Problems*, 37(1), 109–137. <https://doi.org/10.1177/009145091003700106>
- Jones, S. C., Hall, S., & Kypri, K. (2017). Should I drink responsibly, safely or properly? Confusing messages about reducing alcohol-related harm. *PLoS ONE*, 12(9), e0184705. <https://doi.org/10.1371/journal.pone.0184705>
- Jongenelis, M. I., Pratt, I. S., Slevin, T., Chikritzhs, T., Liang, W., & Pettigrew, S. (2018). The effect of chronic disease warning statements on alcohol-related health beliefs and consumption intentions among at-risk drinkers. *Health Education Research*, 33(5), 351–360. <https://doi.org/10.1093/her/cyy025>
- Kaskutas, L. A., & Graves, K. (1994). Relationship between cumulative exposure to health messages and awareness and behavior-related drinking during pregnancy. *American Journal of Health Promotion*, 9(2), 115–124.
- Kaskutas, L. A., & Greenfield, T. K. (1992). First effects of warning labels on alcoholic beverage containers. *Drug and Alcohol Dependence*, 31(1), 1–14.
- Kaskutas, L. A., Greenfield, T., Lee, M. E., & Cote, J. (1998). Reach and effects of health messages On drinking during pregnancy. *Journal of Health Education*, 29(1), 11–17.

- Kersbergen, I., & Field, M. (2017). Alcohol consumers' attention to warning labels and brand information on alcohol packaging: Findings from cross-sectional and experimental studies. *BMC Public Health*, 17(1), 123. <https://doi.org/10.1186/s12889-017-4055-8>
- Krischler, M., & Glock, S. (2015). Alcohol warning labels formulated as questions change alcohol-related outcome expectancies: A pilot study. *Addiction Research & Theory*, 23(4), 343–349. <https://doi.org/10.3109/16066359.2015.1009829>
- Kukla, R. (2010). The ethics and cultural politics of reproductive risk warnings: A case study of California's Proposition 65. *Health, Risk & Society*, 12(4), 323–334. <https://doi.org/10.1080/13698571003789708>
- Laughery, K. R., & Wogalter, M. S. (2016). Designing Effective Warnings. *Reviews of Human Factors and Ergonomics*, 2(1), 241–271. <https://doi.org/10.1177/1557234X0600200109>
- Laughery, K. R., Young, S. L., Vaubel, K. P., & Brelsford Jr, J. W. (1993). The Noticeability of Warnings on Alcoholic Beverage Containers. *Journal of Public Policy & Marketing*, 12(1), 38–56.
- Martin-Moreno, J. M., Harris, M. E., Breda, J., Møller, L., Alfonso-Sanchez, J. L., & Gorgojo, L. (2013). Enhanced labelling on alcoholic drinks: Reviewing the evidence to guide alcohol policy. *European Journal of Public Health*, 23(6), 1082–1087. <https://doi.org/10.1093/eurpub/ckt046>
- Miller, E. R., Ramsey, I. J., Baratiny, G. Y., & Olver, I. N. (2016). Message on a bottle: Are alcohol warning labels about cancer appropriate? *BMC Public Health*, 16, 139. <https://doi.org/10.1186/s12889-016-2812-8>
- Monk, R. L., Westwood, J., Heim, D., & Qureshi, A. W. (2017). The effect of pictorial content on attention levels and alcohol-related beliefs: An eye-tracking study. *Journal of Applied Social Psychology*, 47(3), 158–164. <https://doi.org/10.1111/jasp.12432>
- National Health and Medical Research Council. (2009). *Australian Guidelines to Reduce Health Risks from Drinking Alcohol*.
- O'Brien, P., Gleeson, D., Room, R., & Wilkinson, C. (2018). Commentary on 'Communicating Messages About Drinking': Using the 'Big Legal Guns' to Block Alcohol Health Warning Labels. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 53(3), 333–336. <https://doi.org/10.1093/alcalc/agx124>
- Parackal, S. M., Parackal, M. K., & Harraway, J. A. (2010). Research paper: Warning labels on alcohol containers as a source of information on alcohol consumption in pregnancy among New Zealand women. *International Journal of Drug Policy*, 21, 302–305. <https://doi.org/10.1016/j.drugpo.2009.10.006>
- Peschel, A. O., & Orquin, J. L. (2013). A review of the findings and theories on surface size effects on visual attention. *Frontiers in Psychology*, 4, 902. <https://doi.org/10.3389/fpsyg.2013.00902>
- Peschel, A. O., Orquin, J. L., & Mueller Loose, S. (2019). Increasing consumers' attention capture and food choice through bottom-up effects. *Appetite*, 132, 1–7. <https://doi.org/10.1016/j.appet.2018.09.015>
- Pettigrew, S., Jongenelis, M., Chikritzhs, T., Slevin, T., Glance, D., & Liang, W. (2014). Developing cancer warning statements for alcoholic beverages. *BMC Public Health*, 14(786).
- Pettigrew, S., Jongenelis, M. I., Glance, D., Chikritzhs, T., Pratt, I. S., Slevin, T., . . . Wakefield, M. (2016). The effect of cancer warning statements on alcohol consumption intentions. *Health Education Research*, 31(1), 60–69. <https://doi.org/10.1093/her/cyv067>
- Pham, C., Rundle-Thiele, S., Parkinson, J., & Li, S. (2018). Alcohol warning label awareness and attention: A multi-method study. *Alcohol and Alcoholism*, 53(1), 39–45. <https://doi.org/10.1093/alcalc/agx087>
- Quantum Market Research. (2019). *DrinkWise Australia Pregnancy labelling consumer research. Topline results*.

- Robertson, K., Thyne, M., & Hibbert, S. (2017). Drinkers' perceived negative alcohol-related expectancies: Informing alcohol warning messages. *Drugs: Education, Prevention and Policy*, 24(2), 197–205. <https://doi.org/10.1080/09687637.2016.1188880>
- Rout, J., & Hannan, T. (2016). *Consumer awareness and understanding of alcohol pregnancy warning labels*. Retrieved from Health Promotion Agency website: <https://www.hpa.org.nz/research-library/research-publications/consumer-awareness-and-understanding-of-alcohol-pregnancy-warning-labels>
- Scholes-Balog, K. E., Heerde, J. A., & Hemphill, S. A. (2012). Alcohol warning labels: Unlikely to affect alcohol-related beliefs and behaviours in adolescents. *Australian and New Zealand Journal of Public Health*, 36(6), 524–529. <https://doi.org/10.1111/j.1753-6405.2012.00934.x>
- Siggins Miller. (2014). *Evaluation of the voluntary labelling initiative to place pregnancy health warnings on alcohol products*. Retrieved from <http://foodregulation.gov.au/internet/fr/publishing.nsf/Content/reports>
- Siggins Miller. (2017). *Second evaluation of the voluntary labelling initiative to place pregnancy health warnings on alcohol products: Final report*. Australia. Retrieved from <http://www.health.gov.au/internet/fr/publishing.nsf/Content/pregnancy-warnings-alcohol-labels>
- Stautz, K., & Marteau, T. M. (2016). Viewing alcohol warning advertising reduces urges to drink in young adults: An online experiment. *BMC Public Health*, 16, 530. <https://doi.org/10.1186/s12889-016-3192-9>
- Stockwell, T. (2006). *A Review Of Research Into The Impacts Of Alcohol Warning Labels On Attitudes And Behaviour*.
- Subbaraman, M. S., Thomas, S., Treffers, R., Delucchi, K., Kerr, W. C., Martinez, P., & Roberts, S. C. M. (2018). Associations Between State-Level Policies Regarding Alcohol Use Among Pregnant Women, Adverse Birth Outcomes, and Prenatal Care Utilization: Results from 1972 to 2013 Vital Statistics. *Alcoholism, Clinical and Experimental Research*. Advance online publication. <https://doi.org/10.1111/acer.13804>
- Tam, T. W., & Greenfield, T. K. (2010). Do Alcohol Warning Labels Influence Men's and Women's Attempts to Deter Others from Driving When Intoxicated? *Human Factors and Ergonomics in Manufacturing*, 20(6), 538–546. <https://doi.org/10.1002/hfm.20239>
- Thomson, L. M., Vandenberg, B., & Fitzgerald, J. L. (2012). An exploratory study of drinkers views of health information and warning labels on alcohol containers. *Drug and Alcohol Review*, 31(2), 240–247. <https://doi.org/10.1111/j.1465-3362.2011.00343.x>
- Warren, K. R. (2015). A Review of the History of Attitudes Toward Drinking in Pregnancy. *Alcoholism, Clinical and Experimental Research*, 39(7), 1110–1117. <https://doi.org/10.1111/acer.12757>
- Wigg, S., & Stafford, L. D. (2016). Health Warnings on Alcoholic Beverages: Perceptions of the Health Risks and Intentions towards Alcohol Consumption. *PLoS ONE*, 11(4), e0153027. <https://doi.org/10.1371/journal.pone.0153027>
- Wilkinson, C., Allsop, S., Cail, D., Chikritzhs, T., Daube, M., Kirby, G., & Mattick, R. (2009). *Alcohol Warning Labels: Evidence of impact on alcohol consumption amongst women of childbearing age*. Report 2. Australia. Retrieved from Prepared by National Drug Research Institute (Curtin University of Technology), in collaboration with Drug and Alcohol Office (WA), National Drug and Alcohol Research Centre (University of New South Wales), Public Health Advocacy Institute (Curtin University of Technology) website: <http://www.foodstandards.gov.au/code/applications/Documents/Alcohol-warning-labels-report-2.doc>
- Wogalter, M. S., & Mayhorn, C. B. (2017). Warning design: Signal words. In A. Black, P. Luna, & O. Lund (Eds.), *Information design: Research and practice* (1<sup>st</sup> ed., pp. 338–339). United Kingdom: Taylor & Francis.

- Wogalter, M. S. (2006). Communication-Human Information Processing (C-HIP) Model. In M. S. Wogalter (Ed.), *Human factors and ergonomics. Handbook of Warnings* (pp. 51–61). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Wogalter, M. S., Jarrard, S. W., & Simpson, S. N. (1992). Effects of Warning Signal Words on Consumer-Product Hazard Perceptions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 36(13), 935–939.  
<https://doi.org/10.1177/154193129203601303>
- Wogalter, M. S., Kalsher, M. J., & Rashid, R. (1999). Effect of signal word and source attribution on judgments of warning credibility and compliance likelihood. *International Journal of Industrial Ergonomics*, 24(2), 185–192.
- Wogalter, M. S., & Leonard, S. D. (1999). Attention capture and maintenance. In M. S. Wogalter & D. M. Dejoy (Eds.), *Warnings and risk communication* (pp. 123–148). London: Taylor & Francis.
- Wogalter, M. S., & Rashid, R. (1998). A Border Surrounding a Warning Sign Affects Looking Behavior: A Field Observational Study. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 42(23), 1628.  
<https://doi.org/10.1177/154193129804202357>
- Zahra, D., Monk, R. L., & Corder, E. (2015). 'IF You Drink Alcohol, THEN You Will Get Cancer': Investigating How Reasoning Accuracy Is Affected by Pictorially Presented Graphic Alcohol Warnings. *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 50(5), 608–616.  
<https://doi.org/10.1093/alcalc/agv029>
- Zuckerman, A., & Chaiken, S. (1998). A Heuristic-Systematic Processing Analysis of the Effectiveness of Product Warning Labels. *Psychology & Marketing*, 15(7), 621–642.

## Appendix A: Method

---

### Focus of the review

The key aim of this literature review was to inform the design and content of pregnancy warning labels for packaged alcohol beverages. The focus of the literature review was on the effectiveness of warning labels on packaged alcohol. Effectiveness was conceptualised with reference to Argo and Main's (2004) five dimensions of warning label effectiveness: attention; reading and comprehension; recall; judgement; and behavioural compliance.

Both design factors and the informational content of warning labels influence an individual's response to that warning label. By design factors we are referring to the factors of size; location/placement; colour/contrast; use of signal words; use of pictorials; message length; and physical interactivity. Informational content refers to the actual words and images used to convey information about the risks of alcohol to the individual.

The focus of the literature review was developed by one officer in consultation with the project team.

### Scope

Wilkinson et al (2009) undertook a systematic search for available and relevant literature on the effectiveness of warning labels published from 1989 until October 2008. This covered the period from the introduction of mandatory alcohol warning labels on alcohol in the US (November 1989) to the time of their write up in late 2008. We commenced searches for this review from the end of their search period.

To be included in the review the following inclusion criteria had to be satisfied:

**Study type:** any empirical studies that reported any measures of the five dimensions of effectiveness, including quantitative (e.g. cross-sectional designs, experimental between/within group experimental designs, conjoint designs and choice experiments) and qualitative studies (e.g. in-depth and semi-structured interviews, focus groups). Published reviews were also included.

**Warning type:** text and pictorial based warnings on alcoholic beverages.

**Language:** English only

**Date of publication:** between November 2008 to July 2019

The initial focus was on peer-reviewed published literature. However a considerable grey-literature exists arising from commissioned studies and reports by government and non-government organisations. These were also included where they were empirical studies that reported on the dimensions of effectiveness for populations in Australia and/or New Zealand.

Editorials and commentaries, and other studies that did not report empirical results were excluded from the review.

The scope for the literature review was developed by one officer in consultation with the project team.

## Search

We searched the following on-line electronic databases:

- SocINDEX with full text
- EconLit with Full Text
- Food Science Source
- Food Science and Technology Abstracts
- Medline with full text

The search string used across all databases was : AB ( warning OR advisory OR caution OR label ) AND AB (alcohol)

Using the search software the search was limited to studies that were:

- from peer-reviewed sources
- in English language
- published since November 2008

The search was conducted on 1 August 2019.

Hand searching was also undertaken to identify further papers that may not have been captured through the electronic search. The hand searching included review of reference lists of in scope publications.

Supplementary searches: Two supplementary searches were undertaken to supplement the literature identified regarding signal words given strong interest in the use of terms including 'PREGNANCY ADVICE', and the limited recent evidence on the use of signal words.

Supplementary search 1 was for additional empirical studies on the use of signal words. The search string used across all databases was : AB((warn\* OR risk\* OR threat\* OR danger\* OR "signal word\*" OR caution\* OR care\* OR "take care" OR hazard\* OR "health warn\*" OR "government warn\*" OR alert\*) AND (label\* OR pack\* OR list\* OR claim\* or contain\* OR "consumer good" OR "consumer goods" OR cosmetic\* OR "summary statement" OR "nutrition facts" OR "nutrition information" OR "back of pack" OR "back-of-pack" OR "front of pack" OR "front-of-pack" OR alcohol\* OR beverage\* OR drink\* OR smok\* OR smoking\*)). Results were screened on title and abstract and four additional empirical papers were added to the review.

Supplementary search 2 was to explore the use of the term 'PREGNANCY ADVICE' or 'PREGNANCY WARNING' as a signal word. The search used across all databases was: AB((pregnan\*) AND ( warning OR advisory OR caution OR label OR advice) AND AB (signal word\*). No additional studies were identified.

Searching was undertaken by two officers.

## Screen

The identified papers were up-loaded into EppiReveiwier software and duplicate studies were identified and removed. Fifteen duplicate studies were removed.

Following the removal of duplicates the remaining papers were screened by title and abstract and those that were not in scope were removed. The following types of papers were removed during screening on title and abstract:

- journal editorials
- comment papers
- studies on perceptions and behaviour in relation to standard drinks, %ABV information
- studies on perceptions and behaviour in relation to nutrition and ingredient qualities of alcohol
- studies on advertising alcohol
- clinical and physiological studies of alcohol
- psychological and behavioural studies focussed on treatment.

Full-text versions of in scope studies were then obtained. This set of studies were quality assessed and the relevant information and findings were extracted. Sixteen studies that were out of scope were excluded at the full text review stage (Table A5), leaving 47 studies for quality assessment and data extraction.

Systematic and narrative reviews identified through the search and through hand searching were included within the review

Screening was undertaken by one officer.

## Quality assessment

The relative quality of each study was assessed as being low, medium or high. The quality was assessed with reference to 5 key criteria and two additional criteria for grey literature. The criteria apply across all types of study design. However the criteria were separately operationalised for both qualitative and quantitative studies (Table A1). The quality assessment was derived from a system developed by the US FDA (Verrill & Fanfan 2019).

Each criteria was assessed using a scale of low, medium and high. A study was considered low in the criteria when there was insufficient information to assess, or there were considerable flaws in the study. A study was considered high in a criteria when the paper had full discussion and there were few flaws in the study. A study was considered medium when there was adequate discussion and there were some flaws in the study. Following the assessment of the paper along each criteria, an overall relative assessment of low, medium or high was given to the study.

**Table A1: Quality assessment criteria**

Criteria	Qualitative studies	Quantitative studies
Study aims, justification, & theory	<p>The study is appropriately justified with clear research/study aims.</p> <p>The study is appropriately informed by prior research and theory.</p> <p>Human research ethics committee considerations</p>	<p>The study is appropriately justified with clear research/study aims.</p> <p>The study is appropriately informed by prior research and theory.</p> <p>Human research ethics committee considerations</p>
Study design & methods	<p>The study aims justifies the use of a qualitative approach.</p> <p>Interview and/or facilitation guides are appropriate to study aims.</p>	<p>The study aims justifies the use of a quantitative approach.</p> <p>Measures, instruments, tools are appropriately described and referenced.</p>

Criteria	Qualitative studies	Quantitative studies
		Allocation to any experimental conditions is random.
Population & sampling	<p>Population being recruited is relevant to the aims of the study.</p> <p>Method of recruitment is described, use/how recruitment is stratified is described.</p> <p>Justification for achieved recruitment provided (e.g. saturation).</p> <p>Use of rewards/payments for participation made explicit.</p>	<p>Population being sampled is relevant to the aims of the study.</p> <p>Methods of random sampling is described, how frame is obtained, use/how sampling is stratified.</p> <p>Use of panels is described, how panel is maintained, how members are recruited to panel.</p> <p>A priori sample size and power calculations undertaken, sample size justified.</p> <p>Use of rewards/payments for participation made explicit.</p>
Analysis	<p>Analysis appropriately described and results reported</p> <p>Approach to analysis explained (e.g. grounded theory, thematic coding).</p> <p>Use of video/audio recording and transcription described.</p> <p>Coding approach described, use of a priori coding frame, iteration and definition of codes, multiple coders, measures of inter-coder reliability.</p> <p>Management of researcher bias.</p>	<p>Analysis appropriately described and results reported.</p> <p>Details of any statistical tests used described, and appropriate levels of supporting data reported (test statistic, p-value, degrees of freedom, effect size).</p> <p>Discussion of any assumptions in analysis, appropriate diagnostic test undertaken.</p> <p>Hypotheses testing used where appropriate.</p> <p>Discussion of any weighting undertaken.</p> <p>Confounding variables considered.</p>
Reporting	<p>Reporting of findings, use of participant quotes, negative case reporting.</p> <p>Limitations of approach as discussed.</p>	<p>Reporting of findings, evidentiary support for claims provided.</p> <p>Limitations of approach discussed.</p>
Peer review (for grey literature only)	Studies not peer reviewed were penalised.	Studies not peer reviewed were penalised.
Conflict of interest (for grey literature only)	Studies commissioned or undertaken by non-government organisations with an interest in alcohol warning labels were penalised.	Studies commissioned or undertaken by non-government organisations with an interest in alcohol warning labels were penalised.

Quality assessment was undertaken by one officer.

## **Data extraction**

Data extraction was completed simultaneously with quality assessment. As the paper was read key features and findings were extracted and summarised in an Excel spreadsheet. The five dimensions of effectiveness were used as column headings to organise the data extracted. Additional columns covered study design, sampling, and the quality assessment criteria.

Data extraction was undertaken by one officer.

## **Write-up and synthesis**

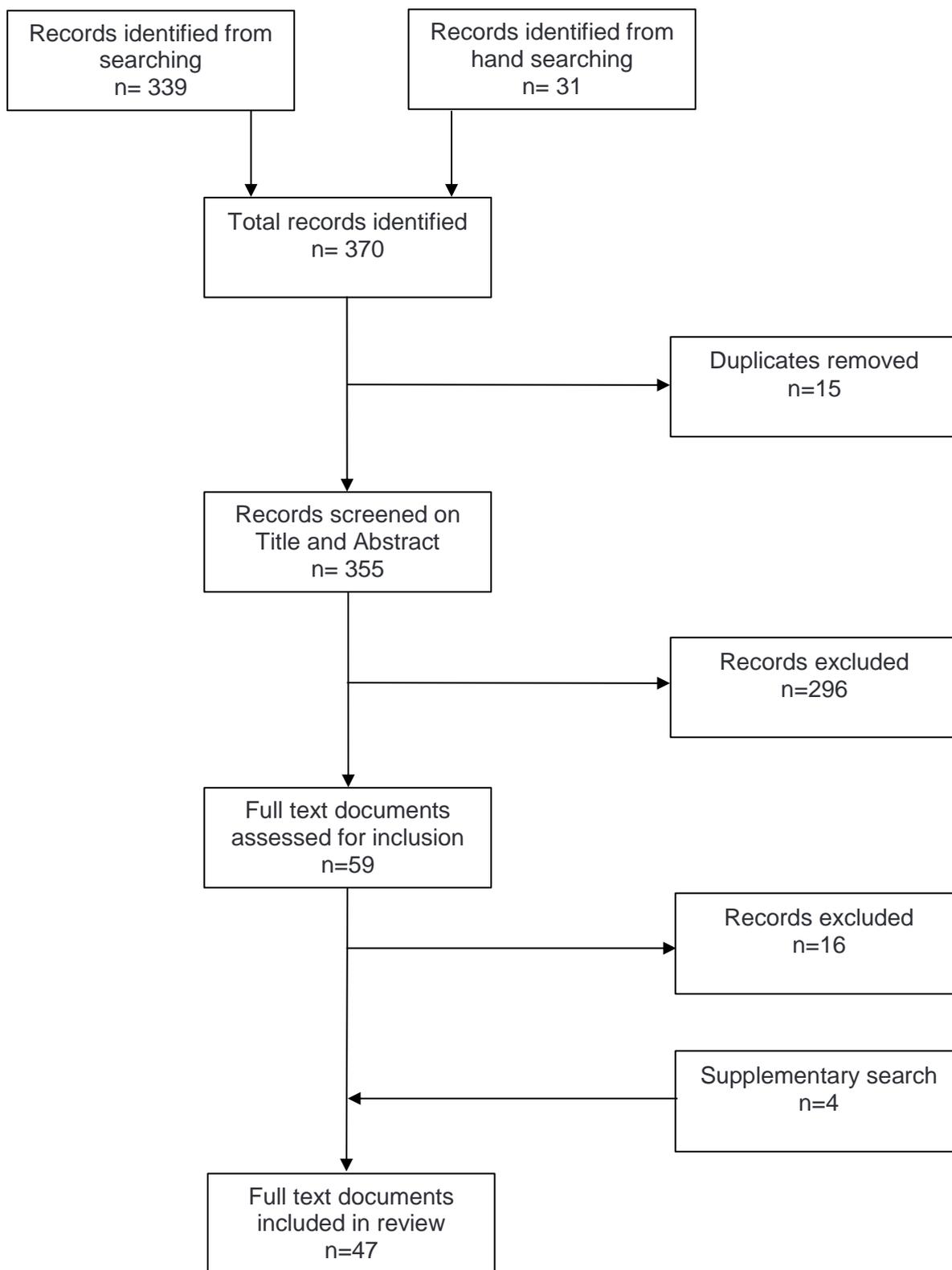
The five dimensions of effectiveness provided the structure for write-up of the findings from each study. Higher quality studies and those that used Australian and/or New Zealand populations were given more weight than lower quality and those that used International populations. The majority of studies reviewed were assessed as being of medium quality. In the write-up, unless otherwise stated studies were assessed as being of medium quality. Where the write-up is about studies that were assessed as low or high quality the quality has been articulated explicitly in the narrative.

Write up and synthesis was carried out by one officer.

## **Peer review**

After completion the draft review was peer reviewed by an independent external academic with expertise in the behavioural sciences. The peer review comments and suggestions were considered and incorporated into the final version of the literature review.

Peer review was carried out by one external academic.



**Figure A1: Literature search flowchart**

**Table A2: Peer-reviewed empirical studies included in the review**

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
Al-hamdani & Smith (2015)	Canada University/hospital adult n=92	Test the impact of warning types and size on consumer and product-based evaluations across beer, wine and spirits.	Quantitative – between subjects  4 levels of warnings: none, text only, combined text and image, combined text and image on plain packaged.  Use real life stimuli cancer warning  Text only was 25% of front package area, combined text and graphic was 50%.	Combined warning and combined warning on plain packaged significantly reduced product and consumer evaluations for all alcohol types except consumer based evaluation on beer.  Text only warning only had significant reduction in product based reductions for spirits – no other significant effects.	Low  small sample convenience sample subject to bias
Al-hamdani & Smith (2017b)	Canada adults consumed alcohol in previous 12 months convenience sample n=440	Test the impact of warning types, size and plain packaging on consumer and product-based evaluations across beer, wine and spirits.	Quantitative – between subjects  Warning size: medium (50% of front label); large (75%); extra-large (90%)  Across branded and plain packed stimuli of beer, wine and spirits  Warning was always text and graphic image	No significant main effect with warning size across product, consumer evaluations and product boringness.  Significant interaction with alcohol type for product based evaluation – larger warnings on spirits and wines had lower product based ratings than smaller warnings on wines and spirits. No significant effect for beer.  As no control (i.e. a no warning product) the base is the 50% sized warning. Authors conclude that	Medium  convenience sample; subject to bias

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
			about alcohol and liver cancer	there is a ceiling effect, after which the gains from increases in size are marginal.	
Annunziata et al (2017)	Italy University students 18-30 years old n=385	Analyse the interest and attitude of Italian university students to health warnings on alcohol and identify segments among consumers.	Quantitative – survey  5 different warnings included (driving, medicine, underage, pregnancy and brain)	Three groups of young consumers with varying degrees of attention to and perceived utility of warning labels. Moderate drinkers view warning labels positively, but less so among riskier drinker.  Higher utility attached to warnings about avoiding drinking if taking medicine and not to drink and drive, with less utility attached to warnings concerned with long terms effects of alcohol such as brain damage and fetal risk.  More emotional impact attached to negative framed warning (wrecked car with fatal message) versus a more generic stylised warning (don't drink and drive and a car logo with strike though).	Medium  small sample, stated preferences for utility and attention – no experimental manipulation.
Annunziata et al (2019)	Italy & France wine consumers Generation Y (born 1978-2000) n=500	Analyse generation Y preferences for and interest in different formats of health warnings.	Quantitative: discrete choice experiment  Attributes in choice set: 1) alcohol content; 2) warning framing; 3)	The warning is the most important element driving utility (61.4%), however it is the presence of the brain warning that negatively influences choice.  Second in contribution to utility (19.3%) is the position of the	Medium  Small sample, obtained sample through social media and word of

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
			<p>warning size; 4) warning location.</p> <p>Warning framing levels: 1) long-term effect (brain damage); 2) short term effect (car crash); 3) no warning.</p> <p>Choice task buying a bottle of wine for dinner with friends. Selecting 1 bottle from a choice of three all costing the same price, all same varietal, both front and back labels of the three bottles shown. Participants to complete 4 choices.</p>	<p>warning where being placed on the front of the label negatively influences choice.</p> <p>Then message with a preference against negatively framed and finally size, with a preference against large warnings. (When analysed by country this second and third swap for Italian sample)</p> <p>Preference for no logo warning option, short term warning and a small logo placed on the back of the product.</p>	<p>mouth potential for bias.</p>
<p>Boluarte et al (2011)</p>	<p>EU-27 Youth 15-24 n=12312</p>	<p>Explore the association between state level policies in the EY-27 and risk perceptions of youth</p>	<p>Quantitative – modelling</p> <p>Secondary data analysis</p>	<p>Significant association between increased alcohol risk perception and two policies: 1) blood alcohol limit for driving; and 2) health warnings on advertisements and/or alcoholic beverage containers</p>	<p>Medium</p>
<p>Coomber et al (2018)</p>	<p>Australia Young adult drinkers 18-25 n=40 in 6 focus groups</p>	<p>Understand young adult drinkers’ perceptions of current voluntary Australian</p>	<p>Qualitative – focus group</p>	<p>Low salience of current voluntary warnings, small size and location on back panel suggest not serious about warning consumers.</p>	<p>Medium</p>

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
		alcohol product warnings.		Current warnings unlikely to influence consumption behaviour, nor encourage them to seek further information from the DrinkWise website, nor encourage discussions with peers or colleagues	
Coomber et al (2017)	Australia University students 18-15 years old n=26	Explore young adult drinkers perceptions of pictorial and graphic warnings	Qualitative – focus group  Using stimuli (pictorial: simplified drawn; graphic: colour photographs) to assess perceptions of emotional, cognitive and intended behavioural responses	<p>Pictorial and graphic warnings stand out to young adult drinkers and convey new information. They elicit a negative emotional response, and may be prompted to reduce the amount they drank, but unlikely to deter them from risky drinking.</p> <p>Large size and front of pack placement increased noticeability, though there was a preference for back of pack position of warnings.</p> <p>Images increase salience. Photos and images led to quick understanding of message without need to read warning text. Photos elicited stronger response than pictorial.</p> <p>Level of unease and discomfort with some warnings – negative emotional response to some graphic warnings Some talk of avoidance through using a glass or</p>	High

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
				turning bottle around to avoid looking at the image  Not stop them drinking but make them think about level of drinking and encourage discussion about drinking with others	
Coomber et al (2017)	Australia Adult drinkers 18-45 n=1061	Investigate awareness of short-term and long-term consequences of alcohol use.	Quantitative – cross sectional survey  Online survey	52.7% aware of warning labels.  Participants aware of any alcohol warning label were significantly more likely than participants not aware of warning labels to respond <i>definitely true</i> to harm to unborn babies and cirrhosis of the liver	Medium
Coomber et al (2015)	Australia Adult 18-45 n=561	Evaluate the awareness of the 'Get the Facts' logo and alcohol warning labels, and evaluate use of the website	Quantitative – cross sectional survey  Online survey	Significant predictors of recalling voluntary warning message were being younger, binge drinker, drink from can/bottle, support warning labels Sex, education, drink type were not significant predictors  similar patterns for other two warnings (is drinking harming yourself or others, kids and alcohol don't mix)	Medium
Critchlow et al (2019)	UK youth and young adult 11-19 years old n=3399	Assess awareness and recall of alcohol health warning information among	Quantitative – cross-sectional survey	13% recalled pregnancy health message which was the second most recalled message (Highest was drink responsibly)	Moderate

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
		youth and young adults	Online survey, weighted to national rep sample	Awareness was significantly higher for those who were of legal drink age, had ever had alcohol, current drinker, high risk drinkers – however some difference may be driven by sociodemographic differences in prevalence of alcohol consumption.	Generalisability limited given youth sample.
Dossou et al (2017)	France Youth 15-29 years old n=26	Assess effectiveness of warning labels across recall, noticeability, credibility, comprehension, responsiveness, impact on behaviour and contrast standard and more embellished packaging.	Qualitative - in-depth interview  Existing products on market with pictogram and print advertisements with mandatory French statement.  Standard and special (celebratory) examples.	Warnings suffered lack of visibility and noticeability due to size and location limitations.  Pictogram generally understood but lacked explicitness.  Source of warning – industry or government a potential credibility risk.  Not likely to encourage behaviour change	Medium  Brief description of methodology and analysis – double coded transcripts.
Dumas et al (2018)	France pregnant or post-partum women n=3603	Assess awareness of the warning and risk perceptions about pregnant and postpartum women	Quantitative – cross sectional	66.1% of women noticed warning label  Drinkers and more educated more likely to be aware of warning.	High
Jarvis & Pettigrew (2013)	Australia Young drinkers	Assess four warning statements for influence on choice behaviour relative to	Quantitative – choice experiment	Brand had the brand highest utility, followed by alcohol content then warning in choice.	Medium  Youth sample – generalisability

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
	18-25 n=300	brand and alcohol content among a youth sample	Qualitative – focus groups  Used realistic brands of RTD beverages varying Brand, alcohol level and warning statement  Warning statement was text, and covered brain health and drink driving. These were identified as most salient in the initial focus groups to test.	5 class latent class model statistically significant  Small influence on some classes - negatively framed messages had higher impact on higher consumers (at risk consumers) positively framed message on driving lead to a positive impact (i.e. increasing utility)	limited to specific populations
Jones & Gregory (2010)	Australia University students 18+ years n=44 in 6 focus groups	Examine the attitudes and opinions of students to warning messages on alcohol and the likelihood of impacts on behaviours	Qualitative – focus groups  Used international products with text based warning statements – including pregnancy warnings.	Some participants had seen warnings before, but thought likely to be ignored.  Not considered effective, not relevant to demographic, or perceived to be targeted to other groups  Limitations in believability	Medium
Jongenelis et al (2018)	Australia Drinkers at long term risk from alcohol 18-65 years n=365	Examine the risk beliefs and impacts of warning statements on those risk beliefs of drinkers who drink at	Quantitative – within subjects experiment  Online survey – simulation of trip from doctor to home via bus stop or vice versa –	Warning labels led to significant increase in the extent to which participants considered alcohol was a risk factor in a related risk (e.g. diabetes, cancer, heart disease and mental illness). There was no significant increase for the liver	Medium  No control group with no warning – but within subjects design offsets this.

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
		long-term risk levels term.	<p>warnings displayed 5 times from 5 different sources in three scenes.</p> <p>5 warnings tested across various outcomes (mental health, cancer, diabetes liver damage, heart disease)</p>	<p>warning though already the highest and about the level others reached following exposure. No moderating effects of age, sex, education or SES.</p> <p>Similar results occurred for drinking intention where the warning reduce intention to drink significantly for all warnings except liver and heart disease warnings. No moderating effects of age, sex, education or SES.</p>	
Kersbergen & Field (2017)	UK University staff and students Study 1 n=60 Study 2 n=120	To explore the attention that warning labels receive.	<p>Quantitative – cross-sectional/eye tracking</p> <p>Quantitative – between subjects design</p> <p>Study 1 was a cross-sectional eye ting study that explore the attention to warnings</p> <p>Study 2 was an between subjects experiment to explore the impact of a motivation intervention on attention to warning labels.</p>	<p>Spent less time looking at health information than brand information, but noted roughly proportional to the sizes of areas.</p> <p>Alcohol warning labels were attended to when they were larger in size and less complex.</p> <p>Manipulation that encouraged participants to focus attention on warning labels did not affect their drinking intentions.</p>	<p>Medium</p> <p>Small sample with potential for bias.</p>

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
Krischler & Glock (2015)	Luxemburg & Germany university students and their colleagues n=122	Investigate the effectiveness of warning labels tailored toward young adults' positive outcome expectancies.	<p>Quantitative - between subjects' design</p> <p>3X2 mixed design with warning labels (questions vs statements vs control) as a between subjects factor and expectancy category (positive vs negative) as within subjects factor</p> <p>3 outcome expectancies presented as warnings with a photographic image: loosen inhibitions (being sick in the toilet); meet new people (police officer); tests your limits (car crash)</p> <p>Presented bottles with the warnings across three bottle of wine and three of beer</p>	<p>Warnings as questions were able to increase individual negative-alcohol related expectancies</p> <p>No impact on individual positive expectancies or general expectancies and drinking behaviour</p> <p>Statement wording warnings had no significant impact on expectancies above control.</p> <p>Authors suggest that the pictures may attract attention and reduce impact of text.</p>	<p>Medium</p> <p>Small sample with potential for bias</p>
Laughery et al (1993)	US Adults Study 1 n=75 Study 2 n=72 Study 3 n=24	Assess the noticeability of warning labels on alcoholic beverages, and testing the influence of various design	<p>Quantitative</p> <p>Study 1 used a set of existing warnings on alcohol that were characterised by their design features and</p>	<p>When container labels were filled with non-warning information locating the warning took longer</p> <p>Pictorials, signal icons and colour (red) can significantly improve</p>	<p>Medium</p> <p>Sample size limit generalisations.</p>

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
		features on noticeability.	<p>recorded the time taken for the participants to correctly identify the warning.</p> <p>Study2 manipulated the presence and absence of 4 design features (colour, pictorial, signal icon and borders) to test the noticeability of the warning</p> <p>Study 3 used eye-tracking to measure time taken to identify warning location.</p>	<p>noticeability. Borders may not have much effect.</p> <p>Presence of a pictorial was the only design feature that significantly lowered response times when tested by eye tracking.</p>	
Miller et al (2016)	Australia Adults n=1547	To test 4 alcohol warning labels about cancer on a variety of impacts (awareness, conversations, drink less often etc)	<p>Quantitative – cross sectional survey</p> <p>Online survey</p>	<p>Females, reading labels and reassessing behaviour based on general product warnings independently predicted agreement that warnings would raise awareness and prompt conversations about cancer risk associated with alcohol consumption, prompt to drink less and prompt to educate children about alcohol-cancer link.</p> <p>Cancer-related warning labels generally perceived as neutral to positive.</p>	Medium

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
				More detailed messages were received more positively.	
Monk et al (2017)	UK University students n=22	Examine attention levels to different types of warnings using eye tracking	Quantitative – between subjects  text and graphic images on white card as stimuli	Participants spent longer looking at graphical images than the text  This was regardless of pictorial content - explicit or not explicit  Pictorial content may cue positive expectancies – caution advised in using graphic imagery	Low  Very small sample, exploratory only
Parackal et al. (2010)	New Zealand teenage and non-pregnant women 16-40 years n=1129	Report the preference of warnings on alcohol as a source of information on risks of alcohol consumption while pregnant.	Quantitative – cross-sectional survey  WATI survey using RDD	Majority of women gave a high rating for preference of warning label as source of information (High 53%; medium 17%; low 30%).  Maori, Pacific and Asian women were more likely to prefer warnings as a source of information than European	High
Pettigrew et al (2016)	Australia Adults consumed alcohol on two days in preceding week 18-65 n=1680	Test the effectiveness of cancer warning statements delivered through a simulation of various situations	Quantitative – within subjects experiment  Participants were exposed to 1 of 6 possible cancer warning statements 3 were general cancer statements and 3 were	Participants were significantly more likely to report a perceived need to reduce alcohol consumption and intention to do so after exposure to the warning.  Generally specific cancer warnings had more enhanced attitudinal outcomes than general cancer warnings.	High

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
			<p>specific cancer statements.</p> <p>Warnings were presented on alcohol beverages containers as well as advertisements, from a child via a health class in school, billboard.</p> <p>In total they saw the same warning in five different situations. Pre and post exposure intentions to drink were recorded.</p>	<p>Participants who found the statement more believable and more personally relevant, reported greater pre – to post-exposure changes.</p> <p>Warning effectiveness was not significantly influenced by participant characteristics.</p>	
<p>Pettigrew et al (2014)</p>	<p>Australia Adult drinkers 18-65 years n=48 in 6 focus groups n=2168 (R)</p>	<p>Develop and test a series of cancer warning statements.</p>	<p>Mixed methods – focus group followed by between subjects survey</p> <p>Developed 12 cancer warning statements varying by message frame (positive/negative), form of cancer (specific/general), fear appeal (fear/no fear) and use of numerical information (yes/no).</p>	<p>Females, younger respondents and those with higher education generally found statements to be more believable, convincing and relevant.</p> <p>Positive framed messages more believable than those with fear appeals and numerical evidence. Similar trend for convincingness and relevance – but some differences by sex</p> <p>Use of 'increase risk' type statements found more believable</p>	<p>High</p>

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
				<p>and convincing than 'can cause' among females</p> <p>More general statements were more believable, convincing and personally relevant than specific statements</p>	
Pham et al (2018)	<p>Australia Adults Study 1 n=559 Convenience sample through advertising and snowballing Study 2 n=87 university students</p>	<p>Assess the impact of changes to size and colour of standard pictogram on attention</p>	<p>Quantitative – between groups experiment Quantitative – eye-tracking</p> <p>Study 1 used an online survey to allocate participants to 1 of four experimental groups 1) control, 2) increased pictogram by 50%; 3) change pictogram from black/grey to black and red and 4) increased size and colour manipulations.</p> <p>Study 2 used eye-tracking to record the time and fixations of participants on the same stimuli as in study 1</p>	<p>Attention was highest for colour and size combination, then size, then colour over control, however effect sizes were small</p> <p>Not all participants looked at warning logo - 65.5% of sample 59% looked at control and 81% looked at increased size and colour warning (a 37% increase with colour and size) However no significant differences in number of fixations, time to first fixation, nor fixation duration average</p>	<p>Medium</p> <p>Small sample size</p> <p>Convenience sample</p>

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
Sillero-Rejon et al (2018)	UK Adult consumed excess alcohol (UK guidelines) in preceding week recruited from university (staff, students and public) 18+ n=128	Study used eye tracking to measure visual attention, with a between-subjects factor of self-affirmation and within subject factor of warning severity to assess defensive and positive reactions to pictorial health warnings.	Quantitative – experiment  On-line large cans of beer - 6 moderately-severe and 6 highly severe occupying bottom 1/3rd of front of branded cans	Self-affirmation did not have any effect on degree of attention  No evidence of effect of warning severity on visual attention	Medium  Convenience sample  Small sample size
Tam and Greenfield (2010)	US adult drinkers n=1376	Develop a conceptual model reflecting the effects of exposure to alcohol warning message an taking actions to avert another’s driving under the influence of alcohol.	Quantitative – cross sectional survey	Recall of the warning message (standardised $\beta = 0.08$ ) had a significant positive direct effect on intervening to stop another’s driving while influenced by alcohol.  Usually consuming 3 or more drinks on a drinking day ( $\beta=0.28$ ) and greater handling of alcohol beverages ( $\beta=0.08$ ) also had significant positive direct effects on intervening.  In addition to the direct effect the relationship between container handling and intervention is mediated by recall of the recall of the warning message.  The findings also indicated the varying effects of gender (women more likely than men), age (40 years and younger more likely than	Medium

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
				older) and ethnicity (ethnic minority more likely than white)	
Thomson et al 2012	Australia teenage and adult drinkers Aged 16 and above n=45 in 6 focus groups	Test the response to a range of warning labels.	Qualitative - focus groups  (Note paper also includes cross-sectional survey but not relevant to this review)	<p>Signal word of 'warning' most acceptable while 'government health warning' was criticised as being akin to 'Nanny' state/Big Brother.</p> <p>Participants generally indicated a preference for new health information in warnings rather than already known information.</p> <p>Graphic images important to increase attention participants linked to experience with tobacco warnings.</p> <p>Simple, clear, unambiguous language in statements was best understood by the greatest breadth of participants. Educate and inform not be authoritarian or prescriptive of behaviour. Serious tone was more acceptable than a humorous tone.</p> <p>Messages needed to resonate with relevant population – the pregnancy warnings did not resonate with males and some young women.</p>	Medium

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
Vallance et al (2018)	Canada Yukon residents (rural/regional population) 19+ years n=45 across 5 focus groups	Explore consumer perception and acceptability of enhanced alcohol labels.	Qualitative – focus group  Used pictorial stimuli of enhanced alcohol labels that included warning message as well as information on low risk drinking guidelines.	Strong support for enhanced alcohol labelling  Participants thought warning about increased risk of cancer were important – considered new information for some participants and linked to a ‘right to know’.  Pregnancy pictogram could be enhanced with additional text about the risk of FASD. Some noted it could aid in discussion with friends and family.  Participants noted that a combination of text and pictogram would likely be more user friendly for people with different levels of literacy.	Medium
Wogalter et al (1992)	US Youth and young adults (high school and college) n=90	Test impact of different signal words on hazard perceptions	Quantitative – within subjects  Five signal words NOTE, CAUTION, WARNING, DANGER, LETHAL were used on mock products with two controls: 1) warning and no signal word 2) no warning and no signal word. Additional test conditions combined an icon	Significant main effect for signal word where all signal words except NOTE were associated with significantly higher hazard perceptions scores.	Medium  Small sample with potential for bias.

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
			(exclamation mark in triangle) next to DANGER and LETHAL		
Wogalter et al (1999)	US University students Study 1 n=66 Study 2 n=57	Testing the effects of presence of signal words and source denoting prefixes on credibility and behavioural compliance	<p>Quantitative – within groups</p> <p>Study 1 tested 6 conditions 1) control no signal word, 2) WARNING, 3) GOVERNMENT WARNING, 4) US GOVERNMENT WARNING 5) FEDERAL GOVERNMENT WARNING 6) US FEDERAL GOVERNMENT WARNING</p> <p>Study 2 tested three groups of sources: specific regulatory agencies (US SURGEON GENERAL's WARNING), professional organisations (e.g. AMERICAN MEDICAL ASSOCIATION</p>	<p>Higher ratings were produced with the signal words presence than its absence and adding more specificity and length.</p> <p>Inclusion of specific sources produced higher ratings compared to signal word alone.</p>	<p>Medium</p> <p>Small sample with potential for bias</p>

Authors	Country, study population and, sampling	Study aim	Design	Key findings	Study quality
			WARNING), general (e.g. HEALTH WARNING)		
Wogalter & Rashid (1998)	US University staff and students n=1200+	Test effectiveness of different borders on the attention a warning receives in a field setting	Quantitative – observational study  Warning with different borders (red thick, yellow/black alternating stripes, thin red, thin black) and controls of warning without borders and no warning at all.	Thick red and thick yellow/black borders were noticed more often and for longer than the control of no border or thin borders.	Low  Abstract only.
Wigg & Stafford (2016)	UK University students 18-35 years old n=60	Test effectiveness of alcohol warning types on risk perceptions and behavioural intentions	Quantitative – between subjects experiment  Pictorial, text-only and no warning on wine and beer bottles.	Pictorial warnings were associated with significantly higher risk perceptions of alcohol and fear arousal compared to control and text-only warnings.  Pictorial warnings were associated with significantly more favourable intention to reduce and quit alcohol consumption.	Medium  Small sample with potential for bias
Zahra et al (2015)	UK University students Study 1 n=153 Study 2 n=58	Investigate the cognitive processing of emotive pictorial warnings intended to curb alcohol misuse	Quantitative – between subjects  Study 1 used Wasson section tasks to test effectiveness of	No difference in reasoning accuracy between positive and negative warning messages.  Accuracy was enhanced when consequences of alcohol were	Medium  Small sample with potential for bias

<b>Authors</b>	<b>Country, study population and, sampling</b>	<b>Study aim</b>	<b>Design</b>	<b>Key findings</b>	<b>Study quality</b>
			negatively valenced graphic warnings.  Study 2 using visual conditionals (if ...then..) based on real word warnings to explore reasoning	considered from negative framed warnings.	

**Table A3: Systematic and narrative reviews included in the review**

<b>Authors</b>	<b>Review type</b>	<b>Content</b>
Argo & Main (2004)	Systematic review	Quantitative studies that reported effect sizes for attention, reading and comprehension, recall, judgement or behavioural compliance across consumer products.
Hassan & Shiu (2018)	Systematic review	Empirical quantitative and qualitative studies on alcohol warning effectiveness
Laughery & Wogalter (2016)	Narrative review	Review of warning studies across consumer products.
Scholes-Balog et al (2012)	Systematic review	Quantitative and qualitative studies on impacts of alcohol warnings with adolescent sample (11-18)
Stockwell (2006)	Narrative review	Quantitative studies on the impact of alcohol warning labels on attitudes and behaviour.
Wogalter & Leonard (1999)	Narrative review	Quantitative studies on warning attention capture and maintenance on consumer products and work place situations
Wogalter (2006)	Narrative review	Detailing a theoretical model and how empirical studies link across consumer product warnings
Wilkinson & Room (2009)	Narrative review	Quantitative and qualitative studies on effectiveness of alcohol warning labelling
Wilkinson et al. (2009)	Narrative review	Quantitative and qualitative studies of on effectiveness of alcohol warning labelling

**Table A4: Grey literature included in the review**

Authors	Country, study population and, sampling	Study aim & Commissioning agency	Design	Key findings	Study quality
GALKAL. (n.d.)	Australia adults 18+ years purchased packaged alcohol in previous 12 months n=301	Measures awareness, understanding and behaviours responding to DrinkWise suite of warnings  Commissioned by DrinkWise Australia	Quantitative: Cross sectional online survey  Stimuli used complete set of DrinkWise warnings	Moderate levels of awareness, high levels of self-reported understanding and some self-reported behaviour change in response to warnings.	Low  Limited detail provided on methodology and sampling. 2+4 slide ppt report.
Hall & Partners. (2018)	Australia adults from Sydney city, suburbs and Newcastle infrequent alcohol consumers screened out  18-45 years n=6-8 in each of 8 focus groups	Investigate consumer understanding and interpretation of the two most commonly used warnings and explore whether there might be potential to enhance effectiveness  Commissioned by Foundation for Alcohol Research and Education.	Qualitative: Focus groups  4 focus groups with women who are pregnant or trying to conceive 2 with female peers 2 with male partners of those trying to conceive  Focus groups were further segmented across age and SES  Used existing warning on products as stimuli as well as new alternative versions of pictograms and warning statements.	Standard pictogram found to be well understood and conveyed message not to drink alcohol while pregnant.  Voluntary warning text seen as ambiguous and weak due to term 'safest'. Reinforces belief that risk from small amounts of alcohol while pregnant is negligible.  Size and location on products of standard pictogram and warning text not attention grabbing due to size.  Other pictograms tested highlighted that: pictogram plus text was more attention grabbing than text alone; red strikethrough seen as universal prohibition symbol and instantly recognisable; size of bump could infer stage of pregnancy; 'harm' is	Medium  No detail on management of moderator and coder bias in data analysis  No coding frame

Authors	Country, study population and, sampling	Study aim & Commissioning agency	Design	Key findings	Study quality
				vague while 'birth defects' was confronting.  Use of personalised language and avoiding definitive statements enhances relevance and credibility.	
Quantum Market Research. (2019)	Australia Adults 18+ female boost sample of 18-44 n=660 (gen pop) n=515 (female boost)	Assess general and target audience awareness and attitudes towards pregnancy labels on alcohol products.  Commissioned by DrinkWise Australia	Quantitative: cross sectional online survey  Stimuli warning with DrinkWise website, pictogram and voluntary warning text.	40% recall in adults (56-59% targets groups) 98% understood message (97-98%) 98% message important and influential 26% took action (28-37%)  Decreasing awareness with age.	Medium  Limited description of methodology and sampling, coding for open ended questions unclear. 3+7 slide ppt report.
Rout & Hannan (2016)	New Zealand Adults booster sample women 18-34 and women with children under 15 n=1,488 (total) n=387 (women 18-34) n=388 (women with children under 15)	Assess the effectiveness of the current warning labels focus on recall, awareness, and reading and comprehension.  Commissioned by Health Promotion Agency	Quantitative: cross sectional online survey  Used standard pictogram, voluntary text and alternate text of 'Don't drink pregnant' as separate stimuli	25% aware of pictogram (30-46% in target populations) 29% aware of voluntary text (36-49%) 19% aware of alternate text (29-20)  22%-62% pictogram prompted various behaviours 38-72% voluntary text prompted various behaviours 23-64% alternative text prompted various behaviours	High  Included discussion of sampling, weighting, quotas used. Coding frames for open ended questions and questionnaire included.
Siggins Miller. (2017)	Australia Adult comprising 6 target groups: 1) pregnant women; 2)	Evaluation of awareness and understanding of	Quantitative cross sectional online survey	38.9% awareness of pictogram among target women 32.5% awareness of text among target women	Medium  Coding frame for open ended

Authors	Country, study population and, sampling	Study aim & Commissioning agency	Design	Key findings	Study quality
Consumer survey in section 4 and Appendix 4	women planning pregnancy; 3) women with child under 18 months; 4) men who is partner of 1, 2 or 3; 5) parent of adult child in 1, 2, or 3; 6) adults over 18 n=5622	warning labels on alcohol  Commissioned by the Australian Department of Health		83.2% understand pictogram as don't drink alcohol when pregnant among target women 1.9% understand pictogram as alcohol causes harm among target women 52.5% understand voluntary statement as don't drink alcohol when pregnant among target women 30.0% understand voluntary statements as alcohol causes harm among target women	questions not provided. Sample boosted by self-enrolment through online banner advertising.

**Table A5: Studies excluded during full text review**

<b>Study</b>	<b>Reason for exclusion</b>	<b>Comment</b>
Al-hamdani (2014)	Commentary	Extended commentary on lessons from tobacco warnings
Al-hamdani and Smith (2017a)	Editorial	Editorial for an edition of <i>Journal of Public Health</i>
Annunziata et al (2016)	Enhanced nutrition label	Primary focus on enhanced nutrition labels on alcohol, limited reference to warnings
Bazzo et al. (2012)	Advertising	Evaluation of an image used in an alcohol and pregnancy advertisement
Brennan et al. (2017)	Off topic	Survey of awareness and attitudes held towards DrinkWise
Brown et al. (2016)	Advertising	Experiment to explore impacts of alcohol promoting and alcohol warning advertising
Cil (2017)	Warning signs	Ecological study on warning signs
Hassan and Shiu (2018b)	Editorial	Editorial for an edition of <i>Alcohol and Alcoholism</i>
Jones et al. (2017)	Off topic	Focus of paper was the interpretation of alcohol industry responsible drinking initiatives
Kukla (2010)	Off topic	Theoretical argument on the ethics and cultural politics of reproductive risk warnings
Martin-Moreno et al. (2013)	Enhanced nutrition label	Primary focus on enhanced nutrition labels on alcohol, limited reference to warnings
O'Brien et al (2018)	Editorial	Response to a commentary
Robertson et al. (2017)	Off topic	Study of negative alcohol related expectancies among a NZ sample
Subbaraman et al. (2018)	Warning signs	Ecological study on warning signs
Stautz and Marteau (2016)	Advertising	On-line experiment on impact of viewing alcohol advertisements
Warren (2015)	Off topic	Historical review of the attitudes toward drinking in pregnancy