



This is a repository copy of *How important is the name in predicting false recognition for lookalike brands?*.

White Rose Research Online URL for this paper:
<http://eprints.whiterose.ac.uk/116655/>

Version: Accepted Version

Article:

Humphreys, M.S., McFarlane, K., Burt, J. et al. (3 more authors) (2017) How important is the name in predicting false recognition for lookalike brands? *Psychology, Public Policy, and Law*, 23 (3). pp. 381-395. ISSN 1076-8971

<https://doi.org/10.1037/law0000133>

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Reuse

Unless indicated otherwise, fulltext items are protected by copyright with all rights reserved. The copyright exception in section 29 of the Copyright, Designs and Patents Act 1988 allows the making of a single copy solely for the purpose of non-commercial research or private study within the limits of fair dealing. The publisher or other rights-holder may allow further reproduction and re-use of this version - refer to the White Rose Research Online record for this item. Where records identify the publisher as the copyright holder, users can verify any specific terms of use on the publisher's website.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>

How important is the name in determining the probability of falsely recognising a lookalike brand?

Michael S. Humphreys, Kimberley A. McFarlane, Jennifer S. Burt, Sarah J. Kelly

The University of Queensland

Kimberlee G. Weatherall

The University of Sydney

Robert G. Burrell

The University of Sheffield; The University of Melbourne

Author Note

Michael S. Humphreys, Kimberley A. McFarlane, and Jennifer S. Burt: School of Psychology, The University of Queensland; Sarah J. Kelly: School of Business, The University of Queensland; Kimberlee G. Weatherall: School of Law, The University of Sydney; Robert G. Burrell: School of Law, The University of Sheffield, and School of Law, The University of Melbourne.

- This research has been supported by Australian Research Council Linkage Grant LP120100249. Financial support for the project came from the Australian Government (via the Australian Research Council), and Australia's Intellectual Property office ("IP Australia"). The Federal Court of Australia, Treasury Wine Estates, and Carlton & United Breweries were non-financial partners who were consulted during the course of the research. Portions of this paper were presented at 'Testing Trade Mark Law's Image of the Consumer', a Meeting and Workshop held at the offices of Carlton & United Breweries, Melbourne, 28 March 2014. This was an invitation only workshop involving lawyers, and external legal and design consultants of the firm. Portions were also presented at 'Testing Trade Mark Law's Image of the Consumer' 20 October 2015. This was a closed workshop attended only by IP Australia employees at IP Australia.

Address correspondence to: Michael Humphreys, School of Psychology, The University of Queensland, St. Lucia, QLD 4072, Australia. Phone: +61 7 3300 8626,

Email: mh@psy.uq.edu.au

Abstract

An under-exploited role for psychology in trademark law is the testing of explicit or implicit judicial assumptions about consumer behaviour. In this paper we examine an assumption that is common across Commonwealth countries, namely, that similar packaging is unlikely to cause consumer confusion provided the brand names are dissimilar. We began by selecting branded products commonly found in supermarkets. For each existing brand we created two novel (fictitious) brands with highly similar packaging to the existing brand. One of these 'lookalike' products had a similar name, the other a dissimilar name. Across two yes/no and one forced-choice experiments using photographs of the real and fictitious products we looked at false recognition rates. Contrary to the judicial assumption participants largely ignored the brand names when making their decisions based on memory. It was only when the pictures of the products were placed side-by-side (in the forced-choice task) that they paid the brand name any significant attention.

Keywords: recognition memory, false memory, trademark law, marketing, advertising

HOW IMPORTANT IS THE NAME IN DETERMINING THE PROBABILITY OF FALSELY RECOGNISING A LOOKALIKE BRAND?

At its core trademark law is a law against consumer confusion. This remains the case despite the move in many jurisdictions to bring additional forms of harm (so-called trademark ‘dilution’) within the scope of trademark protection. Psychology can provide insights relevant to determining what kinds of trader behaviour are most likely to cause confusion. But finding the best way to make use of those insights in legal decision-making is not straightforward. If parties offer evidence of actual consumers’ responses to a mark, it often comes in the form of expensive population sample surveys that aim to be representative of the relevant consumer population or more rarely in the form of experiments. Yet such evidence is frequently rejected or discounted, especially outside the United States (Dinwoodie & Gangjee, 2015; Weatherall, 2017). The reasons this evidence is afforded little (if any) weight are complicated. There are tensions, for example, between courts’ desire for evidence that closely reflects real world situations in which consumers would encounter a mark, and the need to keep costs down. Judges may also hesitate to admit scientific evidence that might constrain their decision-making. Preserving room for judgment is important because a judge may need to resolve conflicting goals: for example, trademark law may need to tolerate some consumer confusion in order to avoid restricting competition, by preventing brand owners monopolising common terms needed to describe products (Dinwoodie & Gangjee, 2015). If psychological research is to have a greater impact on trademark law, we need to identify a role for psychologists’ disciplinary expertise that recognises and responds to these legitimate concerns of the judiciary.

At present (too) much of the focus on the role of psychological evidence in trademark law has been on the value of evidence in particular cases. But psychology can also be useful at a higher level of abstraction. Specifically, we can use experimental and survey methods to

test the assumptions trademark law makes about how consumers process information. For example, we have used subjective responses (a survey method) to test the judicial assumption that the beginning of a word is more important than the ending in determining similarity (Burt, McFarlane, Kelly, Humphreys, Weatherall, & Burrell, in press). The important point here is that we can use the results of a survey that does not attempt to obtain a broadly representative sample or an experiment that does not use a high fidelity simulation of a shopping scenario to test whether an assumption generally holds without purporting to answer whether consumers would or would not be likely to be confused in any given case.

This paper tests assumptions made by the law about the relative importance that consumers assign to brand names as compared to similarity of product appearance or ‘get up’. Our motivation for choosing this as our focus was the (often implicit) assumption in many Commonwealth countries that consumers pay more attention to brand names when making their purchasing decisions than they do to product appearance. This assumption has meant that trademark owners have generally struggled in jurisdictions like Australia and the UK to prevent the sale of ‘lookalike’ products – as long as the brand names are different, the assumption is that consumers will not be confused. The leading Australian case on point is the decision of the High Court of Australia in *Parkdale Custom Built Furniture Pty Ltd v Puxu Pty Ltd* (1982) 149 CLR 191. See also *Dr Martens Australia Pty Ltd v Rivers (Australia) Pty Ltd* (1999) 95 FCR 136 and *Zetco Pty Ltd v Austworld Commodities Pty Ltd (No 2)* [2011] FCA 848. For recent confirmation that the law in the UK is to the same effect see, e.g., *George East Housewares Ltd v Fackelmann Gmbh & Co KG* *Schweppes Ltd v Gibbens* [2016] EWHC 2476 and *The London Taxi Corporation Ltd v Frazer-Nash Research Ltd* [2016] FSR 20. In fact the law on this point has not moved on very far in either jurisdiction since the decision in *Schweppes Ltd v Gibbens* (1903) 22 RPC 601. In that case Lord Halsbury memorably commented “if a person is so careless that he does not look, and

does not... 'treat the label fairly,' but takes the bottle without sufficient consideration and without reading what is written very plainly indeed upon the face of the label on which the trader has placed his own name, then you certainly cannot say he is deceived”.

This assumption is contrary to much of the marketing research on consumer perceptions of brand similarity. While brand name confusion is important in many shopping circumstances, particularly in internet search or shopping list reliance, extant marketing research has emphasised the importance of perceptual attributes in distinguishing products. In particular, the marketing literature has examined the rise of copycat products – products which imitate visual packaging attributes such as colour, shape and lettering while (in general) adopting different product names. That literature tests what kinds of visual cues consumers use to distinguish unknown products, and considers how consumers interpret or respond to the use of similar packaging (Loken & Ward, 1990). There are, for example, studies that look at the impact of imitation on consumer willingness to pay: for example, Steenkamp, Van Heerde, and Geyskens (2010) show that perceptions of qualitative differences between national brands and private labels, and willingness to pay more for national brands, are strongly influenced by packaging design. Other studies demonstrate that consumers have a preference for products that imitate a brand leader (e.g. Warlop & Alba (2004); Aribarg, Arora, Henderson & Kim (2014)). Similarly, Gao, Lim and Tang (2016) use visual stimuli to demonstrate that high resemblance but low quality copycat brands can successfully gain entry to the market.

Another strand of the marketing literature focuses on identifying the kinds of imitation that lead consumers to see packaging as similar, or have more difficulty differentiating the imitating product. Miceli and Pieters (2009) propose and test visual imitation (i.e., attribute- based such as colour, shape and name) and thematic- based copycat (e.g., higher order imitation such as country of origin, benefits or values) similarity

dimensions, again demonstrating the importance of all packaging elements in brand similarity. Satomura, Wedel and Pieters (2014) developed a method and metric to assess the visual similarity or features that make it more difficult to consumers to differentiate between leading and copycat brands when they are presented side-by-side. These results are consistent with the notion that during brief exposures, consumers accurately identify brand images on the basis of mostly coarse visual information (Pieters & Wedel, 2012).

We extend the marketing literature in two ways. First, by looking at memory confusion errors for highly similar packaging but with two different levels of name similarity, we examine the combined effects of packaging and brand name in similarity perceptions, rather than assuming that one is more important than the other. To preview our results, our experimental findings suggest that there is little evidence that dissimilar names prevent memory confusion errors in the context of similar packaging. Memory effects are important in trademark law, which posits, in disputes, a consumer who is seeking a product or trademark they have seen before.

Second, we are most interested in how work in psychology interrelates with the kinds of issues that courts grapple with in trademark disputes. Most of the marketing literature concerns itself with exploring consumer responses that fall outside or are only tangentially related to trademark law's core concern with preventing consumer confusion. The core interests of the marketing literature lie in exploring the full nature and scope of consumer responses. In this paper, in contrast, our starting point is an assumption of fact made by courts.

As will become evident, our findings do not support the legal assumption that motivated our enquiry. This suggests that the assumption needs to be rethought. But this does not mean that legal outcomes need necessarily to be different. There may be some alternative policy basis for allowing the sale of lookalikes: i.e., a good reason to prevent trademark

owners from being given a monopoly over product appearance. For example, there may be a legitimate concern about the gradual crowding out of new market entrants as existing market actors monopolise common colours, shapes that are easy to manufacture and the like (cf. in the United States *Valu Engineering, Inc. v. Rexnord Corporation*, 278 F.3d 1268 (Fed. Cir. 2002) confirming the importance of ‘competitive need’ when determining whether a trademark is ‘functional’ and hence unable to be monopolised via trademark law). These concerns may be sufficiently serious, and the difficulties associated with drawing a line between the legitimate use of common shapes and colours and ‘lookalike’ products may be so pronounced, that the optimal position might be to refuse to extend trademark rights to product appearance: either at all or save in cases where there is some special factor at play. The best outcome, therefore, may or may not be for the law to retain its present approach. However, at the very least, the terms of the debate as to whether the law should intervene to prevent the sale of lookalikes need to be recast, as mentioned in the conclusion.

An introduction to the legal issues and the potential role for psychology

The law that deals with ownership and disputes over elements of branding varies significantly from country to country, but there is enough commonality, enforced by international treaty obligations, to make some general points. All countries that are members of the World Trade Organization (ie, most countries) must enable brand owners to register trademarks under article 15 of the Agreement on Trade-Related Aspects of Intellectual Property Rights (or ‘TRIPS’). Trademarks are signs and symbols (such as brand names, logos, colours or forms of packaging) used to indicate the source of goods or services, and can be registered provided consumers will immediately recognise the symbol as indicating that goods or services come from a particular company, or have been educated to see the symbol as a brand indicator (article 15.1). Registered trademarks are the property of the

trademark owner, who has the exclusive right to use the trademark as an indicator of source, and who can sue to prevent other producers of similar goods or services using the trademark or a mark so similar it is likely to lead to confusion (in the sense of causing consumers to think that the alleged infringing goods come from the same trade source) (TRIPS article 16). In many countries, including the United States and European Union, owners of trademarks also now have broader rights to prevent dilution – that is, the use of similar marks in a way that either tarnishes the brand or in a way that reduces the trademark’s uniqueness or distinctiveness (Martino 1996). Registered marks only provide rights in the country where a trademark has been registered,¹ although there is nothing to stop companies registering their marks wherever they have the business to warrant it.

Most countries also provide legal protection for unregistered aspects of branding through general consumer protection laws aimed at preventing misleading commercial behaviour, and/or specialised legal protection against the use of unregistered marks such as passing off, false advertising and similar legal doctrines (Harms 2012). Thus well-known global brands are protected even in countries where they do not register, as are local firms known to consumers who have not chosen to register their marks. Liability under these legal doctrines tends to depend on the first user of some product name, logo or aspect of ‘get up’ showing (a) that the name, logo or aspect of ‘get up’ has become associated with (distinctive of) the first user; and (b) that its use by an alleged infringer will mislead or confuse consumers with respect to the source of the brand or the existence of a commercial connection between the companies (Harms 2012).

Thus in a majority of countries there are several causes of action aimed at protecting aspects of branding, in which liability turns on proof of (likely) consumer confusion. In the case of unregistered trademarks, the whole gist of the legal action is to give the first user the

¹ Although some regions, such as the European Union, have regional registrations (hence the European Union Trade Mark).

ability to stop conduct in the marketplace by second comers that leads to consumer confusion. In the case of registered trademark law, the role of consumer confusion is more complicated in most countries (the United States being something of an exception in this regard). Trademark owners have exclusive rights, so proving consumer confusion is generally unnecessary if a second comer adopts branding identical to a registered mark on the goods or services for which the owner has registered the mark (TRIPS article 16.1). But if the second comer's marking is not identical, then a trademark owner will need to prove it is confusingly similar.² Potential consumer confusion is also relevant at the registration stage: trademark offices will refuse registration of a trademark that is confusingly similar to one that is already registered for similar goods or services. In some cases, the assessment of confusion is very abstract: since trademarks can be registered before they are used, courts and trademark offices are sometimes not really assessing whether consumers are currently likely to be confused by the second comer's use, but whether they would be confused if they knew and had an imperfect memory of the registered mark (Burrell and Handler 2016, 222-229; Dinwoodie and Gangjee 2015; Weatherall, 2017).

Courts have to determine the likelihood of consumer confusion, and do so on the basis of a range of different types of evidence, although it should be noted that approaches vary between jurisdictions. A judge may receive evidence from marketing experts, and they may use evidence of facts that are thought to indirectly prove that confusion is more or less likely (Beebe, 2006). Courts may also take into account the fame of the first user's mark, the sophistication of the likely consumers, how common elements of the mark are in the particular product category (i.e., whether the words, packaging, colours or images are widely used in relation to the goods or services in question), whether or not the defendant

² The terminology here varies from jurisdiction to jurisdiction: 'confusingly similar' is the American terminology; the European Union asks whether similarity would cause a 'likelihood of confusion'; in Australia courts talk about 'deceptive similarity'.

deliberately set out to confuse consumers, and any evidence that real world consumers have been confused. In addition, courts make assumptions about what sorts of similarity are most likely to be confusing: for example, assumptions about the relative importance of the beginnings, endings, or vowel sounds of words (Burt et al., in press); the relative importance of brand names versus other aspects of packaging; or elements of a product's 'get up' that consumers will remember.

Whether these assumptions are accurate is an empirical question, but is often not treated that way by the courts. In theory, the central place of confusion in trademark law creates an opening for psychology. In the US, consumer surveys have become an accepted way to prove consumer confusion (Cohen, 1991), and psychologists have often been involved in the design of the large-scale representative surveys submitted to courts. In the UK and Australia, however, courts are sceptical about the benefits of empirical evidence designed to test likely consumer confusion (Weatherall, 2017). Courts in these jurisdictions require parties to prove the value of any surveys or experiments before they are done. Only studies that meet a high standard of validity and reliability will be admitted, which necessitates an expensive process of ensuring a representative sample,³ and careful question design. In the case of ordinary consumer goods, courts frequently either refuse to admit empirical evidence (on the basis that it is unnecessary) or allow it in but disregard or discount the results. Often, a single judge will end up weighing up all the various facts and making their own judgment as to whether they think confusion is sufficiently likely to warrant the court's intervention – often making empirical assumptions in the process (Weatherall, 2017).

We do not take the view that psychology should displace legal judgment. Deciding whether a defendant has crossed the line between permissible and impermissible similarity

³ There is a greater reliance on selecting samples with "relevant" demographic characteristics than from randomly sampling from a defined population. This is especially true when online samples are used.

must be made in a broader context that takes into account not merely competition concerns, but also issues of what might be called legal policy – how the parties have conducted themselves during the litigation, whether a case has been properly pleaded and so forth. Finding the proper role for psychology is therefore far from straightforward. We would, however, suggest that empirical results can at least ensure that judges' reasoning (assumptions) regarding how consumers think and act are based on reality. Questions of memory, in particular, are important in the legal analysis of confusion. The process of assessing likely confusion involves considering the response of a consumer who has an 'imperfect recollection' of the claimant's trademark or product packaging and then sees (e.g., on a supermarket shelf or grocery website), or hears (e.g., a radio advertisement or shopping instructions from a partner) the defendant's brand. One step towards convincing judges of the value of psychological methods generally is to undertake studies that are designed to test common judicial assumptions about consumer responses.

The present work seeks to test one of the assumptions that courts make either explicitly or implicitly in cases involving lookalike products in Commonwealth countries. Namely, that confusion is unlikely as long as the brand or product name is different (Burrell & Handler, 2016). In the Australian *Maltesers* decision, for example, an imitator adopted packaging very similar to the leading chocolate brand, using the same bright red colour on packaging that depicted spherical chocolate balls, some of them sliced in half to show a yellowish centre. The competitor's product, however, was called 'Malt Balls'. It was held that given the difference in name, and the likely perfect memory consumers would have for the famous brand, consumer confusion was unlikely (*Mars Australia Pty Ltd v Sweet Rewards Pty Ltd* (2009) 81 IPR 354).

Psychological Issues

The assumption that shoppers will not be confused by similar packaging if the brand name is different is an assumption about how people recognize stimuli in context. Deciding on the validity of the assumption is complex, given that brand name similarity and packaging similarity are multivariate concepts. Furthermore, understanding what shoppers attend to turns on questions about how existing and novel names are processed, the type of comparison process allowed by the shopping scenario (simultaneous or successive comparisons) and what the customer is looking to purchase.⁴

In order to produce meaningful results, we therefore sought to simplify the factors at play. In our experiments the novel (imitator) brands were all fictitious so that we could control for prior knowledge. We selected well-known supermarket brands and then created two sets of fictitious brand names. Most of the fictitious names in the first set were designed to be similar in orthography to a well-known brand, i.e. they shared multiple letters in common (Edgell/Vegell). A small proportion of the first set were selected for similarity of meaning (Green's/Emerald's). The second set of names were dissimilar in both orthography and meaning (Duracell/Supaunit). A complete list of names is provided in Appendix A. Inevitably there is scope to challenge our assessments of 'similarity' since there was an unavoidable degree of subjectivity in determining that two words were similar, and no doubt our word pairings reflect different degrees of similarity (some are almost identical, others are similar but nowhere near identical). Nevertheless, the examples we have chosen are ones where we think there is a good argument that, put to the test, most if not all of the marks would be considered confusingly similar.

⁴ For additional information about the relevance of the work in psychology on recognition in context to trademark law see (Humphreys, McFarlane, Burt, Kelly, Weatherall, & Burrell, 2016).

We used high quality full color screenshots of the packaging (e.g., a can of beans, a breakfast cereal box, etc.). With the similar names the same high quality image of the known brand was used. With the dissimilar names a minor change in the image was made in that we also manipulated the the color and/or font of the name.⁵ This produces a strong match between the contents of memory for the picture and the distractor picture tested, but is unlikely to have any further impact on the results from the yes/no task used in Experiments 1A and 1B. That is, it would have been possible to produce an equally strong match by shortening the retention interval or presenting the studied image for a longer period of time. These manipulations of match strength are possible because it is established that in general people do not have detailed memories for the features of highly familiar objects. For example, there is a considerable amount of evidence showing that Americans do not have a detailed memory for whether the head faces left or right on the Lincoln penny (Kosslyn & Rabin, 1999; Nickerson & Adams, 1979). Of course, the similarity of the pictures is obvious in a forced choice test where a picture of the existing brand is paired with the same or highly similar picture of the novel brand. This change in the saliency of using the same or a highly similar picture for the novel brand in going from a yes/no to a forced-choice test provided much of the motivation for using a forced-choice test in Experiment 2.

Across experiments we looked at simultaneous versus successive judgments to reflect different shopping scenarios. The more common scenario is that the shopper has set out to purchase an existing brand that they have purchased before. In store, they may then encounter a brand with highly similar packaging but with a novel name which may be either similar or dissimilar to the name of the brand that they intend to purchase. The second scenario is that the shopper has decided to purchase a brand for the first time (e.g., a new type of shampoo)

⁵ These images were created for a slightly different purpose but the difference in the similarity of the images was retained in this study. However, this confound does not change our conclusions because it can only enhance the apparent effect of name similarity not reduce it.

and has been told the name of the product and shown the bottle (e.g. by a friend or in an advertisement). Once in store, they encounter a product in packaging that is similar to the brand they are hoping to try. Again, the name used on the 'lookalike' product may be similar or dissimilar to the name of the brand which they encounter in the store. Our goal in this research was to identify the main variables that would impact on brand choice in these scenarios. The legal assumption is that in situations where the rest of the context (the packaging) is very similar, a dissimilar name will generally be sufficient to avoid consumer confusion.

We must account, however, for other psychological processes that may complicate the effects of item similarity. In particular, there may be a tendency to choose the existing name and avoid the novel name. Such a tendency has been reported in the marketing literature (Hoyer & Brown, 1990; Park & Lennon, 2009). The cognitive literature shows that words that occur commonly in text (high frequency words) are more easily identified than lower frequency words (Monsell, 1991). But when shoppers are performing a recognition memory judgment – that is, deciding whether a brand is the same as one that they encountered previously in a specific context (e.g., in a shopping list, or in the study phase of the experiment) – the effects can be complex. Memory research with words shows that episodic recognition of words from a prior study list is inferior for high frequency compared to low frequency words (Glanzer & Bowles, 1976). The latter are less likely to be encountered in everyday life just prior to the experiment, so that confusion among the different encounters with a word is reduced. However, this effect reverses with very low frequency words, which suffer a disadvantage in recognition (Chalmers & Humphreys, 1998; Osth & Dennis, 2014). In our test (reflecting a common phenomenon in consumer marketing), our novel brands are made up although they sometimes contain one or more known words. Due to these

complications, no clear prediction for the level of recognition of existing brands versus our invented brands is possible.

Our research design sought to test the effects of item similarity while accounting for the impact of brand familiarity using a complex within subject (within list) design, in which we used pictures of existing brands with a visible brand name, and highly similar (or same) pictures with novel brand names. The novel names were either similar or dissimilar to an existing brand in the experiment. After studying a subset of brands containing both existing and novel brands, the participant made a recognition decision on three different types of items defined by their relationship to a studied item. There were studied items (targets), related distractors, and unrelated distractors. Related distractors were highly similar or identical product pictures to studied items (not included at test), but had different brand names from the studied items: if the studied item had an existing brand name, the corresponding related distractor had a novel brand name (either similar or dissimilar); if the studied item had a novel brand name, the corresponding related distractor had an existing brand name. Unrelated distractors were novel and existing brands whose product categories had not been included in the study list. Across different study lists, each item on the test list served equally often as a target, a related distractor, and an unrelated distractor.

One possible criticism of this experimental design is that it is highly artificial, and that the task we are asking of our subjects is too far divorced from the task that faces consumers in real world shopping scenarios. In particular, we asked the same subject to study both real, known products (with an existing brand name and packaging), and manipulated products (with essentially the same packaging as an existing but unstudied brand, but an unknown brand name) and tested the same subject on both existing and novel brands. This might seem too different from a consumer who has a (real) memory for an existing product and is searching only for that product when she is confronted at the store with similar-looking

products. However, the kind of confusion we are testing for would be encountered in the real world and would be a concern for trademark law. A consumer could be seeking a product based on only a fleeting encounter with the packaging – via advertising, or being shown an unfamiliar product by a friend or family member. He or she might also be searching for a variety of products, some of which he or she knows from previous shopping experiences and others of which were previously unfamiliar to him or her. Our design tests whether very different names will, in such scenarios, lead to substantially less error.

A complex design of this kind is not common in marketing research, but has several advantages for our purposes. The use of multiple instances of each type of item not only increases the power of the design, but also enhances generality. It also captures the fact that recognition decisions are made against a background of memories for the other items in the list or episode, as well as memories from previous encounters with that stimulus or other stimuli (Dennis & Humphreys, 2001; Humphreys, Bain, & Pike, 1989; Osth & Dennis, 2014). This design enables us to assess whether there is a tendency to choose existing brands over novel brands by comparing the false alarm rate (FAR) for unrelated distractors that are existing brands versus novel brands. It also gives us information as to whether existing brands are better learned than novel brands by comparing the hit rate (HR) for targets that are existing brands versus novel brands. Finally, and most importantly for present purposes, it can tell us how much people are confused by the related distractors through comparing the FARs to related and unrelated distractors. This is done separately for distractors that are existing brands versus those that are novel brands. In other words, we can investigate the complex interactions between brand familiarity, product category, packaging and brand name similarity by looking at different FARs in different relations.

In order to further tease out these interactions in a range of shopping contexts we have used both a yes/no task and a forced-choice task. While forced-choice tasks have been

favoured in marketing research, both types of scenarios occur in shopping contexts. Where original and imitator brands are presented on the shelf side-by-side, a consumer's decision mirrors a forced-choice task. There are, however, shopping scenarios involving presentation of only one product on the shelf, which elicit a yes/no decision process: for example, where a product has temporarily sold out, the products have been placed at a physical distance from one another, or simply because the store only stocks one product. This last scenario is of particular importance in the case of supermarkets that rely heavily on imitation strategies and where only imitator brands may be available in some product categories. Equally importantly, it is the scenario that is most often posited by trademark law, where we tend to be concerned primarily about the danger that a consumer may be confused when encountering the defendant's product in isolation (as witnessed by the law's concern for the possible impact of faulty memory or 'imperfect recollection').

In our yes/no task (Experiments 1A and 1B), participants were asked to remember shopping items representing novel and existing brands that were presented, one at a time, pictorially in their packaging. They were tested in an old/new recognition test on items they had studied, unrelated unstudied items from unseen product categories, and related distractors that were similar in name and packaging to studied items (Experiment 1A) or dissimilar in name but similar in packaging to studied items (Experiment 1B). In each experiment there were two groups which differed in the sampling of the distractors for the test. For the existing brand FAR group, the related distractors were existing brands, whereas for the novel brand FAR group, the related distractors were novel brands.

The design of Experiments 1A and 1B allowed us to answer three questions:

1. The first was our core research question: whether the use of a dissimilar brand name would prevent confusion between the studied brand and the distractor brand with highly similar packaging. If the legal assumption holds true, we

would expect a large difference between the FAR for related distractors in Experiments 1A and 1B (that is, the related FAR would be much higher in Experiment 1A (with similar names) than 1B (with dissimilar names)). We would also expect that the FAR to related distractors in Experiment 1A (similar names) would be substantially greater than the FAR for unrelated distractors, whereas the difference would be small in Experiment 1B (dissimilar names).

2. The second was the impact of other factors: in particular, whether there would be a tendency to choose an existing brand over a novel brand. If this were true, we would have expected our participants to make more false alarms when the unrelated distractor was an existing brand than when it was a novel brand.
3. The final question was whether a novel brand (i.e., new entrant to category) was more likely to be confused for an existing brand (i.e., established category brand) or whether an existing brand was more likely to be confused for a novel brand. Most research has examined the impacts on an existing brand of a novel imitation brand entering the category, but there are shopping scenarios which might produce confusion the other way – perhaps when there is an absence of the novel brand (sell out or supermarket shelf placement strategy), or a person inexperienced with the category is relying upon memory to purchase a brand (e.g., husband shopping for wife’s shampoo). Few studies have examined the potential cross effects of imitation brands and original brands, and whether these effects might be asymmetric. Lawyers are most likely to be concerned about novel brands being mistaken for existing brands, but if the tendency observed in the marketing literature holds true, the opposite form of confusion

may be more likely (existing brand mistaken for studied novel brand) (van Horen & Pieters, 2012). Indeed, the long proven marketing strategy of brand extension relies on this notion that both the original and novel extension brand will benefit from the extension (Broniarczyk & Alba, 1994; Park, Milberg & Lawson, 1991).

Adding Experiment 2, a forced-choice task, enables a number of further insights. By asking participants to choose between pictures of two products shown side-by-side, we can compare simultaneous versus sequential presentation of products to understand whether they result in different levels of confusion. Both lawyers and psychologists would expect better discriminability when products are presented side-by-side, although we would still expect less discriminability where a related distractor with a very similar name is presented, compared to situations where the related distractor has a very different name. Discriminability should also be better with same category pairs than with products from different categories: with highly similar distractors, accuracy is better when the choice is between a target and a distractor that is highly similar to that target (Heathcote, Freeman, Etherington, Tonkin, & Bora, 2009; Tulving, 1981).⁶

EXPERIMENT 1A

Method

Participants

⁶ In the psychology literature, the standard explanation for subjects' better performance where distractors are highly similar is that the "memory strengths" of a target and its highly similar distractor will be positively correlated, with the result that the difference between the two memory strengths will be less variable than when the choice is between a target and a distractor whose memory strengths are uncorrelated. However, in the situation which we will be investigating, this explanation may not apply: in a side-by-side presentation we would expect participants to focus on the only aspects which differentiate the products – that is, the brand names.

Sixty undergraduate students from The University of Queensland participated in Experiment 1A in exchange for credit toward a first-year psychology course. Half of the students were randomly assigned to the existing brand FAR condition, and half were randomly assigned to the novel brand FAR condition. All participants were required to be Australian citizens or permanent residents for four or more years.

Materials

The 36 well-known brands and the pictures of their packaging came from 36 different product categories commonly found in supermarkets. All images were a consistent size (300 x 300 pixels), presented in colour on a white background, and brand names (the well-known name, the fictitious similar name, or the fictitious dissimilar name) were always clearly visible on the packaging. Care was taken to select well-known brands that are only used on a relatively limited range of products, and that the novel brands were not in use commercially in relation to the tested products. See Appendix A for the complete list of product categories with corresponding brand names employed in the experiment.

Design

For Experiment 1A, only the existing and similar-name novel brands were used. The composition of the test list was manipulated between-subjects (existing brand FAR vs. novel brand FAR). In both conditions, half of the study list comprised 12 existing brand items and half comprised 12 novel brand items. In order to create the study lists 24 brand categories were randomly chosen for each participant with half of the categories contributing an existing brand and half contributing a novel brand. In the existing brand FAR condition, the test list consisted of all of the existing brand items that had been studied earlier and half of the studied novel brand items. The other half of the novel brand items were replaced with related distractors. That is, for example, the unstudied existing brand, Duracell was tested after studying Evacell. In the novel brand FAR condition, the test list consisted of all of the

studied novel brand items plus half of the studied existing brand items. The other half of the studied existing brand items were replaced with related distractors. That is, the existing brand, Duracell, was studied and Evacell was tested. In addition, there were six existing brands and six novel brands from 12 unstudied product categories included in both the existing brand and the novel brand FAR conditions. These were the unrelated distractors. Thus in all test lists there were 18 targets and 18 distractors (see Figure 1).

Insert Figure 1 about here

Procedure

The experiment was computer administered, and participants were tested individually. All participants were given instructions prior to the start of the experiment that they would be presented with a “shopping list” of 24 items, with each item presented one at a time on the screen for four seconds each. They were instructed that their task was to memorise the items, as their memory would be tested for the items at a later time. Consistent with this, the study phase of the experiment consisted of study item images being presented in the center of the screen on a white background one at a time for 4000 ms each. Following the study phase, a self-paced lexical decision task unrelated to the experiment was presented that participants took between 15 and 20 minutes to complete. This task was followed by a self-paced single item recognition test.

Prior to the test phase, participants were instructed that the procedure for the single item recognition test would involve the test item images being presented one at a time in the center of the screen with buttons labelled ‘YES’ and ‘NO’ appearing directly underneath each test item image. If participants believed they had seen the item during the study phase, they were required to click on the button labeled ‘YES’. If they believed they had not seen the

item during the study phase, they were required to click on the button labeled ‘NO’. No information was provided about the nature of the distractors. Once a button was clicked, the next test item image would appear. Upon completion of the test phase participants were debriefed.

Results

The mean probability of responding ‘YES’ (old) for existing and novel brands at test is shown in Table 1.

 Insert Table 1 about here

A 2 x 2 mixed-model ANOVA was conducted on HRs with brand type of tested item (existing vs. novel) included as the within-subjects factor and test list condition (existing brand FAR vs. novel brand FAR) as the between-subjects factor. A main effect of brand type was observed, $F(1, 58) = 4.49$, $MSE = .01$, $p = .038$, $\eta_p^2 = .07$, with lower HRs observed for existing brands ($M = .88$) than novel brands ($M = .93$). No main effect of test list condition or interaction between brand type and test list condition were observed, both $F_s < 1$. An independent-groups t-test conducted on the related brand FARs revealed no significant difference between test list conditions, $t < 1$. In both test list conditions, we subtracted the unrelated FAR from the corresponding related FAR for each participant and tested whether these difference scores significantly differed from zero. In the novel brand FAR condition, this difference score (.64) significantly differed from zero, $t(29) = 11.55$, $p < .001$, $d = 2.11$.⁷

⁷ When we have a concern about the magnitude of an effect (difference) we report Cohen's d if we have used a t-test and Cohen's f if we have used an ANOVA (Cohen, 1988). For our one-sample t-tests, Cohen's d is the mean of the difference scores minus the null hypothesis mean (i.e., zero for difference scores) divided by the standard deviation. Cohen proposed d values of .20, .50, and .80 correspond to small, medium, and large effects in the behavioural sciences. Cohen's f is the standard deviation of population means divided by their common

This difference (.67) was also significant in the existing brand FAR condition, $t(29) = 12.47$, $p < .001$, $d = 2.27$.

An additional 2 x 2 mixed-model ANOVA was conducted on the unrelated FARs with brand type of tested item (existing vs. novel) as the within-subjects factor and test list condition (existing brand FAR vs. novel brand FAR) as the between-subjects factor. No main effects or interaction between the factors were observed, all $F_s < 1$. Thus there was no indication that there was a bias to choose an existing brand.

Discussion

The FAR to related distractors was much larger than the FAR to unrelated distractors in both the existing brand FAR condition and in the novel brand FAR condition, but we could not tell whether this was due to the similarity of the packaging and/or the similarity of the brand names. The results of Experiment 1B below, which tested dissimilar names, helps answer this question. The HR was slightly higher for novel brand targets than for existing brand targets. We had no prediction about this comparison because non-words and unknown words are generally recognized less well than known words, and we expected that the real words used in some of the existing brand names would function like known words. However, the use of one or more known words in the novel brand names complicates this prediction. It did not matter whether participants had studied an existing brand and were tested on a highly similar brand, or whether they had studied a novel brand and were tested on a highly similar existing brand as the related FARs in these two conditions were almost identical. The analysis of the unrelated FARs showed that there was no evidence of a tendency to choose an existing brand over a novel brand.

standard deviation; f was calculated using η_p^2 for relevant main effects in these analyses. Cohen proposed f values of .10, .25, and .40 correspond to small, medium, and large effects in the behavioural sciences.

EXPERIMENT 1B

Method

Participants

Sixty undergraduate students from The University of Queensland participated in Experiment 1B in exchange for credit toward a first-year psychology course. None of the participants in Experiment 1A were used. Half of the students were randomly assigned to the existing brand FAR condition, and half were randomly assigned to the novel brand FAR condition. All participants were required to be Australian citizens or permanent residents for four or more years.

Design, Materials, and Procedure

The design, materials and procedure of Experiment 1B were the same as Experiment 1A. The only difference was that novel brands with dissimilar names were used.

Results

The mean probability of responding ‘YES’ (old) for existing and novel brands at test is shown in Table 2.

Insert Table 2 about here

A 2 x 2 mixed-model ANOVA was conducted on HRs with brand type of tested item (existing vs. novel) included as the within-subjects factor and test list condition (existing brand FAR vs. novel brand FAR) as the between-subjects factor. No main effects of brand type, $F < 1$, test list condition, $F(1, 58) = 3.04$, $p = .087$, or interaction between the factors were observed, $F < 1$. An independent-groups t-test conducted on the related FARs for

existing and novel brands revealed no significant difference between test list conditions, $t(58) = 1.64$, $p = .106$.

We subtracted the unrelated FAR from the corresponding related FAR for each participant in both test list conditions and tested whether these difference scores significantly differed from zero. In the novel brand FAR condition, this difference score (.56) significantly differed from zero, $t(29) = 10.44$, $p < .001$, $d = 1.90$.⁸ This difference (.59) was also significant in the existing brand FAR condition, $t(29) = 10.44$, $p < .001$, $d = 1.60$. In order to compare unrelated FARs to existing and novel brands, we subtracted the unrelated FAR to novel brands from the unrelated FAR to existing brands in both the novel brand FAR condition and the existing brand FAR condition. The difference scores in both the novel brand FAR condition (.03) and in the existing brand FAR condition (.08) were significantly different from zero, $t(29) = 2.69$, $p = .012$, $d = .49$, and $t(29) = 2.39$, $p = .024$, $d = .44$, respectively.

A 2 x 2 between-subjects ANOVA conducted on FARs to related distractors with experiment (1A and 1B) and test list condition (existing brand FAR vs. novel brand FAR) included as factors revealed no significant main effect of experiment, $F(1, 116) = 1.85$, $p = .176$, main effect of test list condition, $F(1, 116) = 1.56$, $p = .214$, or interaction, $F(1, 116) = 1.07$, $p = .303$. Effect sizes (Cohen's f) for the main effect of experiment and test list condition were .13 and .11, respectively, corresponding to small effect sizes under Cohen's (1988) guidelines. If the true effect sizes were medium ($f = .25$) or large ($f = .40$), post hoc power analyses using G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) revealed that we would have power of .78 or .99, respectively, to detect significant effects with our current sample of 120 participants.

⁸ In the novel brand FAR condition, no participant made a false alarm to an unrelated distractor when the distractor was a novel brand. This also occurs in the two unrelated FAR comparisons. The logic of testing whether a difference score differs from zero is unaffected.

Discussion

When we subtracted the unrelated FAR from the corresponding related FAR in Experiments 1A and 1B and tested to see if the resulting distributions had means greater than zero, we obtained highly significant results with large effect sizes for both experiments. However, the p values are somewhat lower and the effect sizes are somewhat larger in Experiment 1A where the brand names for the related distractors were similar to a studied brand name than in Experiment 1B where the brand names are dissimilar. Nevertheless, the statistical evidence available does not indicate a sizable reduction in confusion from Experiment 1A to 1B. When we compared the related FARs in Experiment 1A (similar names) with the related FARs in Experiment 1B (dissimilar names) the difference was not significant. Post hoc power analyses showed that we had a .78 or .99 probability of finding a significant effect ($p < .05$) if the true effect size (Cohen's f) was medium (.25) or large (.40). Although it is an arbitrary designation, these effect sizes are considered standard guides in the psychological literature (Cohen, 1988).

Additionally, it is noteworthy that when we compared related and unrelated FARs we observed highly significant results and very large effect sizes in Experiment 1B where dissimilar names were used. For legal purposes, the (large) absolute size of the related FAR with dissimilar names, is also likely to be informative. We defer until after the results from Experiment 2 whether we have grounds for concluding that we have found a counterexample to the generalization that dissimilar names will protect against confusion.

In Experiment 1B, significantly more false alarms were made to unrelated distractors with an existing brand name than to unrelated distractors with a novel brand name. This is consistent with findings from the marketing literature, but is inconsistent with the results from Experiment 1A. A possible explanation for this discrepancy is that the novel brands in

Experiment 1A were not thought of as novel brands. Instead, the combination of the packaging and the similar brand name was perceived to be the existing brand. Unlike Experiment 1A, there was no difference in the HRs between existing and novel brand names. Our conclusion is that any differences are small and largely uninterpretable due to the different kinds of existing and made up names, and the possibility that a bias to choose existing brands has inflated the HRs for existing brands.

EXPERIMENT 2

In Experiment 2 we used a forced-choice test. The primary purpose was to see whether yes/no and forced choice tests produced the same results. In addition, we manipulated the relationship between the distractor and the target in the test pairs in order to learn more about the similar and dissimilar names. Specifically, in some test pairs the study item was paired with an unstudied brand from the same product category, whereas in other test pairs the study item was paired with an unstudied brand from a different product category. However, the unstudied brand from a different product category was similar to another studied brand. When the unstudied brand in a test pair was from the same product category as the studied member of the test pair, we expected participants to ignore the packaging and focus on the name. This should have improved recognition performance. However, we could not predict the likely impact with any degree of certainty because we did not know either the magnitude of any tendency to choose the known brand in a pair, or how much difficulty would be produced by asking participants to choose between similar names.

Method

Participants

Fifty-eight undergraduate students from The University of Queensland participated in Experiment 2 in exchange for credit toward a first-year psychology course. Half of the students were randomly assigned to the similar brand condition, and half were randomly assigned to the dissimilar brand condition. All participants were required to be Australian citizens or permanent residents for four or more years. An additional 31 participants distributed evenly across both conditions (17 from the dissimilar brand condition and 14 from the similar brand condition) were tested, but their data had to be excluded due to a computer fault. Because of the random nature of this exclusion, this does not effect any of our conclusions.

Materials

The materials were the same as those used in the preceding experiments with the exception that all 36 product categories provided study items: 18 categories provided existing brand items for study, while the remaining 18 provided either similar or dissimilar novel brand name items for study depending on the brand type condition the participant was assigned to. The study list therefore consisted of 36 items.

Design and Procedure

The brand type composition of both study and test lists was manipulated between-subjects (similar brand name vs. dissimilar brand name). In the similar brand name condition, half of the 36 item study list comprised existing brand names (e.g., Duracell) and half comprised similar novel brand names (e.g., Evacell). In the dissimilar brand condition, half of the study list comprised existing brand items (e.g., Duracell) and half comprised dissimilar novel brand items (e.g., Supaunit).

At test, 24 test pairs containing one studied item (target) and one unstudied item (distractor) were presented in a forced-choice test. Each test pair consisted of one existing and one novel brand. All of the distractors shared their product category with a target. In

half of the test pairs, the distractor came from the same category as the target (same category condition – e.g., Duracell (target) vs. Evacell (similar) or Supaunit (dissimilar)). In the other half of the test pairs, the distractor came from a studied category other than the target category (different category condition – e.g., Duracell (target) vs. Bright Wings (similar) or Baker’s Help (dissimilar) when both Duracell and White Wings had been studied; see Figure 2). In addition, in half of the pairs the target was an existing brand item and in the other half it was a novel brand item.

Insert Figure 2 about here

The result was a 2 x 2 x 2 factorial design with the presence of similar and dissimilar brand items in the test list manipulated between-subjects, and the composition of the test pairs (same category or different category) and the type of target (an existing brand item or a novel brand item) manipulated within-subjects. The results were analysed as two 2 x 2 ANOVAs: one for test pairs with existing brand items as targets, and one for test pairs with novel brand items as targets.

The procedure was the same as that used in the preceding experiments, except that a self-paced forced choice test was employed instead of a single item recognition test. Prior to the test phase, participants were instructed that the procedure for the forced choice test would involve two test item images being presented together in the center of the screen. Participants were required to click on the item they believed they had seen during the study phase. Once an image was clicked, the next pair of test images would appear. The assignment of targets to the left or right position of the pair was determined randomly for each test pair.

Results

The mean probability of correctly selecting the studied item (target) for existing, novel similar, and novel dissimilar brands at test is shown in Table 3.

 Insert Table 3 about here

To examine existing brand targets, a 2 x 2 mixed-model ANOVA was conducted on probability correct with pair type composition (same category vs. different category) included as the within-subjects factor and novel brand condition (similar to existing brand name vs. dissimilar to existing brand name) as the between-subjects factor. A main effect of pair type was observed, $F(1, 56) = 16.51$, $MSE = .02$, $p < .001$, $\eta_p^2 = .23$, with lower HRs observed for studied existing brands paired with non-studied novel brand items from a different product category ($M = .81$) than for studied existing brands paired with non-studied novel brand items from the same product category ($M = .91$). No main effect of novel brand condition was observed, $F < 1$, and the effect size was very small ($f = .04$). If the true size of this effect was medium ($f = .25$) or large ($f = .40$), post hoc power analyses using G*Power 3.1 revealed that we would have power of .61 or .95, respectively, to detect significant effects with our sample of 58 participants. No interaction between pair type and novel brand condition was observed, $F < 1$.

To examine novel brand targets, a 2 x 2 mixed-model ANOVA was conducted on probability correct with pair type composition (same category vs. different category) included as the within-subjects factor and novel brand condition (similar to existing brand name vs. dissimilar to existing brand name) as the between-subjects factor. Although no main effect of pair type was observed, $F(1, 56) = 3.15$, $p = .082$, a significant main effect of novel brand condition was observed, $F(1, 56) = 4.83$, $MSE = .04$, $p = .032$, $\eta_p^2 = .08$, with lower probability correct observed for studied novel target brands that were similar to existing

brands ($M = .75$) than for studied novel target brands that were dissimilar to existing brands ($M = .83$). The effect size (Cohen's f) for this comparison was $.29$. No interaction between pair type and novel brand condition was observed, $F < 1$.

Two additional 2×2 within-subjects ANOVAs were also conducted on probability correct for each similarity condition separately, with pair type composition (same category vs. different category) and target type (existing vs. novel) included as factors. In the analysis on the similar novel brand condition, although no main effect of pair type was observed, $F < 1$, a significant main effect of target type was observed, $F(1, 28) = 10.33$, $MSE = .04$, $p = .003$, $\eta_p^2 = .27$, with lower probability correct observed for studied novel brands ($M = .74$) than for studied existing brands ($M = .87$). The interaction between pair type and target type was also significant, $F(1, 28) = 8.66$, $MSE = .02$, $p = .006$, $\eta_p^2 = .24$, with simple effects tests for pair type revealing that studied existing brands were better recognised when presented in same category pairs ($M = .92$) than different category pairs ($M = .81$), $F(1, 28) = 9.61$, $p = .004$, $\eta_p^2 = .26$, while studied novel brands showed no significant effect of pair type, with a trend in the opposite direction favoring different category pairs ($M = .77$) over same category pairs ($M = .72$), $F(1, 28) = 1.19$, $p = .285$.

Similarly, the analysis on the dissimilar brand condition also obtained a significant interaction between pair type and target type, $F(1, 28) = 7.61$, $MSE = .02$, $p = .010$, $\eta_p^2 = .21$, with simple effects tests for pair type revealing that studied existing brands were better recognised when presented in same category pairs ($M = .90$) than different category pairs ($M = .81$), $F(1, 28) = 6.92$, $p = .014$, $\eta_p^2 = .20$, while studied novel brands only showed a trend toward being better recognised when presented in different category pairs ($M = .86$) than same category pairs ($M = .80$), $F(1, 28) = 2.26$, $p = .144$. No main effect of pair type or target type was observed, both $F_s < 1$.

Discussion

As expected, when studied existing brands were tested side-by-side with a distractor, performance was better with same category pairs than with different category pairs. There was, however, no indication that the similarity between the target and distractor names in the pair made a difference. A post hoc power analysis indicated that we had probabilities of .61 and .95 to detect a significant effect ($p < .05$) if a medium or large effect size existed. This result is consistent with results showing that, by themselves, the effects of orthographic, phonological, and semantic similarity on false recognitions are generally small (Criss, 2006; Underwood & Humphreys, 1979). This particular result does not support the judicial assumption that, even when contextual elements such as packaging are highly similar, dissimilar names will prevent confusion.

When studied novel brands were tested, those novel brands were less well recognized when paired with similar compared to dissimilar brands. However, because this only happened when the studied brand is novel, the results do not support the judicial assumption about lookalike brands not being confused with existing brands if the name is different. In addition, it is arguable that the difference in the FARs of .08 with a medium effect size of .29 is not so large that we would expect that courts or lawyers would consider it important. There was also a trend for performance to be better with different category than same category pairs. This trend was not significant, but when the results from the similar and dissimilar novel brand conditions were analysed separately there was a significant interaction between pair type (same category vs. different category) and target type (existing brand vs. novel brand) in both conditions.

We believe that some, or all, of this interaction may be due to the tendency noted in the marketing literature to choose the existing brand over the novel brand. We had observed such inclinations in Experiment 1B where participants made more false alarms to unrelated

brands when they were existing brands than when they were dissimilar novel brands. This phenomenon was not observed in Experiment 1A where the novel brands were similar to existing brands. However, with same category forced choice tests it would have been apparent that one brand name was an existing brand and the other was a novel brand. In turn, this may have caused participants to notice that one of the brands in a different category pair is also novel. Although we think that bias processes are almost certainly playing a role, in order to explain the observed pattern of results the bias would have to differ as a function of pair type (same vs. different category), target type (existing vs. novel brand), and novel brand condition (similar vs. dissimilar). The complexity of our findings illustrates the problem with forced-choice tests in that there are too many unmeasured factors influencing performance.

General Discussion: what do these results tell us about trademark law and its assumptions?

Although the present work addresses only some facets of brand recognition by consumers, it is possible to draw some conclusions that have potentially important implications.

First, our results suggest that the use of a similar brand name together with similar packaging will cause people to mistakenly identify a novel brand as an existing brand. This is hardly surprising as a matter of common sense or from the perspective of existing legal doctrine, but it is worth noting that, in this respect at least, trademark law appears to rest on a firm foundation.

Secondly, yes/no and forced-choice recognition tasks seem to produce different results. Providing participants with an opportunity to compare brands side-by-side does cause them to examine names more closely, at least when the packaging is very similar.⁹ This too is

⁹ Notably this is underlined by the fact that with the forced choice tasks, there was a difference between same category and different category pairs. This presumably occurs

consistent with trademark law, which in general concerns itself more with sequential presentation of brands. Insofar as the law is concerned with simultaneous presentation, the assumption tends to be that confusion is less likely to arise. Again, this result is not surprising, but is nevertheless still worth noting – there is value in identifying areas where the law’s assumptions are borne out, not merely where they are misplaced.

Where our experimental findings depart from the law’s assumptions, and hence are more controversial, is in our observation that when a consumer commits a product to memory, his or her recollection is likely to focus in significant part on the ‘get up’ of the product. Consequently, a dissimilar name will not necessarily protect against false brand recognition.

Here it is important to bear in mind a qualification as to what our study seeks to show. It will be remembered that our methodology in Experiments 1A, 1B, and 2 turned on participants being asked whether they had been shown an item during the experiment. This is a reasonable proxy for the type of scenario that attracts much legal attention; namely, the reaction of a consumer who has had some prior exposure to a brand, but that exposure is sufficiently limited such that his or her memory may be hazy or imperfect. For example, someone may have sampled and enjoyed a brand of beer at a friend’s house and formed the view that they would like to purchase a few bottles for themselves if and when they see it for sale. In this scenario a consumer may try to commit the product to memory. Once they are in the store, they may recall that they want to purchase the beer they sampled but have relatively poor memory for the brand name or the packaging. They may attempt to recognize one of the items on the shelf as the one they intend to purchase. Under these conditions the match between the package on the shelf and the memory for the packaging may dominate over any mismatch between the names.

because with same category pairs, participants tried to ignore the packaging and concentrated on trying to determine whether they recognised the name.

Our study does not, however, purport to speak to the scenario where a consumer is very familiar with a brand, such that his or her recollection is much more likely to be fixed. Australian courts have held that where a brand is very well-known, confusion may be less likely because strong prior familiarity will displace the likelihood of imperfect recollection. Interestingly, this sort of reasoning has not found favour in other jurisdictions. Whether Australian courts are right to adopt this approach is precisely the sort of matter on which psychologists ought to have something useful to contribute. But this was not our focus and for this reason alone one could not conclude from our study that the *Maltesers* case discussed in the introduction was wrongly decided.

We also did not attempt to create a high fidelity simulation of our shopping scenarios. These considerations might matter if the intention were to provide evidence in a particular case. Our purpose, however, was to test an assumption (often implicit) in Commonwealth countries that dissimilar brand names will protect against confusion. Unless otherwise specified, such an assumption must be assumed to apply in general to marketing scenarios and to non-marketing scenarios as well (e.g., performance in a psychology laboratory experiment). Our results show that this assumption does not apply in general. This means that it would be hazardous to rely on this assumption in deciding a particular case unless reasons could be articulated as to why it would apply in that case and not in our scenarios. Of course a newly articulated version of this assumption is possible. For example, one might argue that in real world shopping scenarios consumers have learned to rely on brand names because so much of the packaging they encounter is confusable, or that brand names have assumed greater importance because internet searches rely on brand names. One way of testing whether these alternative versions of the assumption are plausible would be to rerun Experiments 1A and 1B while challenging the participants in the test instructions to ignore the packaging and concentrate on the name.

Another implication of our results relates to the way that lawyers and courts approach survey or experimental design. It is not uncommon for surveys or experiments tendered in court to adopt a single design (one survey with one question design, administered to many participants): even though trademark law recognises both confusion that could result from side-by-side and from sequential presentation (assuming confusion is more likely in the latter scenario). With simple stimuli (e.g., words), forced choice and yes/no tasks typically produce very similar results when an appropriate measure of the ability to discriminate between studied and non-studied items is used.¹⁰ However, as our results show, with complex stimuli that have different levels of similarity, an observer probably selects or weights different aspects of the stimulus in a side-by-side comparison versus a sequential comparison. As we theorised above, consumers presented with an existing and imitator brand side-by-side may be forced to focus on the aspects that are dissimilar – here, the brand name. In this situation appropriate discriminability measures are likely to differ between the yes/no and the forced-choice tasks. Because yes/no and forced-choice tasks have different strengths and weaknesses it is frequently necessary to use both kinds of tasks in attempting to understand a difficult problem such as the influence of similarity on recognition.

Finally, our research design also tested for evidence of a tendency noted in the marketing literature to choose familiar brands over less familiar brands. We found some evidence of a tendency to choose familiar brands, but the pattern of results varied over name similarity conditions. When distractors had similar names to the existing brand (Experiment 1A) there were small and non-significant differences between the FARs to unrelated distractors with familiar-looking packages/novel names and unrelated distractors with familiar packages/existing names. However, the same comparisons were substantially greater

¹⁰ Theoretically d' (a measure of discriminability from signal detection theory) calculated from a forced-choice task equals the square root of 2 times d' calculated from a yes/no task. Thus there is more information when a target and distractor are simultaneously presented even though the recognition process is fundamentally the same.

when the unrelated distractors had dissimilar brand names (Experiment 1B). Specifically, participants made significantly more FARs to the existing names. This is a phenomenon based on the pre-existing familiarity of the brands because, by definition, the unrelated distractors were not similar to any brand that had been studied. The failure to find this same effect when the distractor names were similar could suggest that the participants did not perceive these stimuli as being unfamiliar. In other words, the participants appear to be activating aspects of the memorial representation of the existing brand without recognizing the differences in the names. This kind of confusion also occurs in the Moses illusion where people respond “two” to the question “How many animals of each kind did Moses take aboard the ark?” without realizing that it was Noah not Moses (Erickson & Mattson, 1981). There are also models of recognition where the presentation of a similar distractor can cause the memory system to converge either to the long term representation of the target or to the long term representation of the distractor (Chappell & Humphreys, 1994). In the later case the participant would be highly confident that the distractor had been studied.

Generally, there were modest differences between groups (existing vs. novel brand targets at test). The overall picture is that people do not appear to have a much better episodic memory for an existing familiar brand name than they do for a novel unfamiliar brand name. In Experiment 1A when distractors had a similar name to a studied item, the probability of correctly recognizing a studied novel brand was slightly but significantly higher than the probability of correctly recognizing a studied existing brand. In Experiment 1B, there was no significant difference between these measures. In Experiment 2, pairs with an existing brand as the target were better recognized than pairs with a novel brand as the target in the similar name brand condition, but there were no main effects and only an interaction between pair type and target type in the dissimilar name brand condition. However, to some extent the interaction in the dissimilar brand name condition and the main

effect in the similar name brand condition may have been driven by a bias to choose an existing name over a novel name. In summary, the differences in the ability to recognize studied existing and novel brands appear to be small.

Conclusions

At the outset we noted that, in Commonwealth countries and beyond, courts and lawyers have struggled to identify the appropriate role for psychological evidence in trademark proceedings. Surveys and experiments directed to particular legal controversies and tendered as evidence are expensive and, inevitably, subject to thorough criticism that often leads to them being ignored or rejected. The experiments we report here represent an attempt to take a different approach: examining not a particular controversy but a common assumption made by courts about how consumers behave.

The results reported in this paper do support some assumptions lawyers make: that sequential presentation of a brand is more likely to lead to confusion than presentation side-by-side, for example. But we did not find support for the assumption that different brand names will remove the risk of confusion. As we noted in the introduction, this doesn't necessarily mean courts should be reaching different results – there may be good policy reasons to allow imitation of product get up – but it does mean that we should be thinking about the issue in a different way: perhaps by recognising explicitly the policies reflected in our construction of consumer confusion.

Our results also suggest another important point: that attempting to test confusion using psychological methods is complex. Depending on the method used, we observed different results. This suggests that if there is a desire to make better use of psychology in actual legal proceedings, it would be worth exploring whether a more nuanced approach should be adopted: instead of big, representative, single-approach studies, multiple smaller

studies that elucidate responses in a range of scenarios will provide a more accurate picture of what is going on in consumers' minds.

References

- Aribarg, A., Arora, N., Henderson, T., & Kim, Y. (2014). Private label imitation of a national brand: Implications for consumer choice and law. *Journal of Marketing Research*, 51(6), 657-675. doi:10.1509/jmr.13.0420
- Beebe, B. (2006). An empirical study of the multifactor tests for trademark infringement. *California Law Review*, 94(6), 1581-1654. doi:10.15779/Z38571T
- Broniarczyk, S. M., & Alba, J. W. (1994). The importance of the brand in brand extension. *Journal of marketing research*, 31(2), 214-228. doi:10.2307/3152195
- Burrell, R. G., & Handler, M. (2016). *Australian Trade Mark law* (2nd ed.). Melbourne: Oxford University Press.
- Burt, J. S., McFarlane, K. A., Kelly, S. J., Humphreys, M. S., Weatherall, K., & Burrell, R. G. (in press). Brandname confusion: Subjective and objective measures of orthographic similarity. *Journal of Experimental Psychology: Applied*.
- Chalmers, K. A., & Humphreys, M. S. (1998). Role of generalized and episode specific memories in the word frequency effect in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(3), 610-632. doi:10.1037/0278-7393.24.3.610
- Chappell, M. & Humphreys, M. S. (1994). An auto-associative neural network for sparse representations: Analysis and application to models of recognition and cued recall. *Psychological Review*, 101(1), 103-128.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, D. (1991). Trademark strategy revisited. *The Journal of Marketing*, 46-59.

- Criss, A. H. (2006). The consequences of differentiation in episodic memory: Similarity and the strength based mirror effect. *Journal of Memory and Language*, 55(4), 461-478. doi:10.1016/j.jml.2006.08.003
- Dennis, S. & Humphreys, M. S. (2001). A Context Noise Model of Episodic Recognition Memory. *Psychological Review*. 108, 452-478.
- Dinwoodie, G., & Gangjee, D. S. (2015). The image of the consumer in European trade mark law. In D. Leczykiewicz & S. Weatherill (Eds.), *The image(s) of the consumer in EU law*. Oxford: Hart Publishing.
- Erickson, T. A., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, 20, 540-552. doi:10.1016/S0022-5371(81)90165-1
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. doi:10.3758/BRM.41.4.1149
- Gao, S.Y., Lim, W.S., & Tang, C.S. (2016). Entry of copycats of luxury brands. *Marketing Science, Articles in Advance*, 1-18. doi:10.1287/mksc.2016.1008
- Glanzer, M., & Bowles, N. (1976). Analysis of the word frequency effect in recognition memory. *Journal of Experimental Psychology: Human Learning and Memory*, 2(1), 21-31.
- Harms, LTC (2012) *The Enforcement of Intellectual Property Rights: A Case Book* (3rd ed.) Geneva: World Intellectual Property Organization (WIPO)
- Heathcote, A., Freeman, E., Etherington, J., Tonkin, J., & Bora, B. (2009). A dissociation between similarity effects in episodic face recognition. *Psychonomic Bulletin & Review*, 16(5), 824-831. doi:10.3758/PBR.16.5.824

- Horen, F. V., & Pieters, R. (2012). When high-similarity copycats lose and moderate-similarity copycats gain: The impact of comparative evaluation. *Journal of Marketing Research*, 49(1), 83-91. doi:10.1509/jmr.08.0405
- Hoyer, W. D., & Brown, S. P. (1990). Effects of brand awareness on choice for a common, repeat-purchase product. *Journal of Consumer Research*, 17(2), 141-148.
- Humphreys, M.S., Bain, J.D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic and procedural tasks. *Psychological Review*, 96, 208-233.
- Humphreys, M.S., McFarlane, K.A., Burt, J.S. Kelly, S. J. Weatherall, K. G. Burrell, R. G. (2017). *Psychonomic Bulletin & Review*, doi:10.3758/s13423-017-123
- Kosslyn, S. M., & Rabin, C. (1999). The representation of left–right orientation: A dissociation between imagery and perceptual recognition. *Visual Cognition*, 6(5), 497-508. doi:10.1080/135062899394911
- Loken, B., & Ward, J. (1990). Alternative approaches to understanding the determinants of typicality. *Journal of Consumer Research*, 17(3), 111-126.
- Martino, T. (1996) *Trade Mark Dilution*. Oxford: Clarendon.
- Miceli, N., & Pieters, R. (2010). Looking more or less alike: Determinants of perceived visual similarity between copycat and leading brands. *Journal of Business Research*, 63(11), 1121–28. doi:10.1016/j.jbusres.2009.10.007
- Monsell, S. (1991). The nature and locus of word frequency effects in reading. In D. Besner & G. W. Humphreys (Eds.), *Basic Processes in Reading: Visual Word Recognition* (pp. 148-197). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, 11(3), 287-307. doi:10.1016/0010-0285(79)90013-6

- Osth, A. F. & Dennis, S. (2015). Sources of interference in item and associative recognition memory. *Psychological Review*, 122(2), 260-311. doi:10.1037/a0038692
- Park, C. W., Milberg, S., & Lawson, R. (1991). Evaluation of brand extensions: The role of product feature similarity and brand concept consistency. *Journal of consumer research*, 18(2), 185-193.
- Pieters, R., Wedel. M. (2004) Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects. *Journal of Marketing*: April 2004, Vol. 68, No. 2, pp. 36-50. doi: <http://dx.doi.org/10.1509/jmkg.68.2.36.27794>
- Satamura, T., Wedel, M., & Pieters, R. (2014). Copy alert: A method and metric to detect visual copycat brands. *Journal of Marketing Research*, 51(1), 1–13.
doi:10.1509/jmr.11.0467
- Steenkamp, J.E.M., van Heerde, H.J. & Geyskens, I. (2010). What makes consumers willing to pay a price premium for national brands over private labels? *Journal of Marketing Research*, 47(6), 1011-1024.
- Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning & Verbal Behavior*, 20(5), 479-496. doi:10.1016/S002-5371(81)90129-8
- Underwood, B. J., & Humphreys, M. S. (1979). Context change and the role of meaning in word recognition. *The American Journal of Psychology*, 92(4), 577-609.
doi:10.2307/1421791
- Van Horen, F., & Pieters, R. (2012). Consumer evaluation of copycat brands: The effect of imitation type. *International Journal of Marketing*, 29(3), 246-255.
doi:10.1016/j.ijresmar.2012.04.001
- Warlop, L. & Alba, J. W. (2004). *Journal of Consumer Psychology*, 14, 21-27.

Weatherall, K. (2017). The consumer as the empirical measure of trade mark law. *Modern Law Review*, 80(1), 53-83. doi:10.1111/1468-2230.12242

Table 1

Hit Rates (HRs) and False Alarm Rates (FARs) as a function of brand types of tested items and their relationship with studied items for participants in the existing brand FAR and novel brand FAR conditions of Experiment 1A. Related distractors had similar names.

Test List Condition	Brand Type of Test Item	Relationship to studied target		
		Same (HR)	Related (FAR)	Unrelated (FAR)
Existing	Existing	.91	.69	.02
Brand FAR	Novel	.93	--	.03
Novel Brand	Existing	.87	--	.03
FAR	Novel	.92	.68	.04

Note. Both groups studied both novel and existing brands, but the study items targeted at test were novel brands for one group and existing for the other. A related distractor was from the target product category and was an existing brand for a novel target from the study phase, and novel for an existing target. Unrelated distractors were novel and existing brands from unstudied product categories.

Table 2

Hit Rates (HRs) and False Alarm Rates (FARs) as a function of brand types of tested items and their relationship to studied items for participants in the existing brand FAR and novel brand FAR conditions of Experiment 1B. Related distractors had dissimilar names.

Test List Condition	Brand Type of Tested Item	Relationship to studied target		
		Same (HR)	Related (FAR)	Unrelated (FAR)
Existing	Existing	.93	.67	.08
Brand FAR	Novel	.96	--	.00
Novel Brand	Existing	.95	--	.03
FAR	Novel	.94	.56	.00

Note. Both groups studied both novel and existing brands, but the study items targeted at test were novel brands for one group and existing for the other. A related distractor was from the target product category and was an existing brand for a novel target from the study phase, and novel for an existing target. Unrelated distractors were novel and existing brands from unstudied product categories.

Table 3

Probability correct as a function of pair type composition and brand type of studied item for participants in the similar brand and dissimilar brand conditions of Experiment 2.

Novel Brand Condition	Pair Type Composition	Brand Type of Studied Item in Pair		
		Existing	Novel Similar	Novel Dissimilar
Similar to Existing Brand Name	Same	.92	.72	--
	Category			
	Different	.81	.77	--
	Category			
Dissimilar to Existing Brand Name	Same	.90	--	.80
	Category			
	Different	.81	--	.86
	Category			

Study Phase	Test Phase	<i>Existing Brand FAR Condition</i>	<i>Novel Brand FAR Condition</i>
12 existing brand items	Targets	12 existing brand items	6 existing brand items
12 novel brand items		6 novel brand items	12 novel brand items
	Related Distractors	6 existing brand items from the same product categories as the 6 studied novel brand items that were not used as targets	6 novel brand items from the same product categories as the 6 studied existing brand items that were not used as targets
	Unrelated Distractors	6 existing brand items	6 existing brand items
		6 novel brand items	6 novel brand items

Figure 1. A graphical depiction of the study and test phase compositions for the existing brand FAR and similar novel brand FAR conditions in Experiment 1A.

Study Phase	Test Phase		
18 existing brand items	6 x Pair Type: Same Category	Targets	Existing brand items
		Distractors	Similar/dissimilar novel brand items from the <i>same product category</i> as targets
18 similar/dissimilar novel brand items	6 x Pair Type: Same Category	Targets	Similar/dissimilar novel brand items
		Distractors	Existing brand items from the <i>same product category</i> to targets
	6 x Pair Type: Different Category	Targets	Existing brand items
		Distractors	Similar/dissimilar novel brand items from a <i>different product category</i> to targets
	6 x Pair Type: Different Category	Targets	Similar/dissimilar novel brand items
		Distractors	Existing brand items from a <i>different product category</i> to targets

Figure 2. A graphical depiction of the study and test list compositions to illustrate the pair types employed at test in Experiment 2.

APPENDIX A

Table A1

Thirty-six product categories commonly found in supermarkets with names of corresponding known brands, similar brands, and dissimilar brands that were used as stimuli

Product Category	Known Brand	Similar Brand	Dissimilar Brand
Baby Wipes	Huggies	HugMe	Dryhaps
Batteries	Duracell	Evacell	Supaunit
Baked Beans	Heinz	Heanz	Bertie's
Biscuit Snacks	Shapes	Squares	Tasteez
Bread	Bürgen	Berjin	Grainie
Butter	Western Star	Northern Star	Five Points
Cake Mix	Green's	Emerald's	Mix Ups
Canned Vegetables	Edgell	Vegell	Sideveg
Cereal	Weet-Bix	Wheat-Bits	Brekkioes
Cheese	Bega	Bage	Chesi
Crisps	Smith's	Smithies	Crunchysnax
Coffee	Moccona	Maccono	Instacaf
Deodorant	Rexona	Rexano	Temptis
Dishwashing Liquid	Morning Fresh	Dawn Fresh	Ultra Shine

Energy Drink	V	B	Up And Atem
Flavoured Milk	Breaka	Brekka	Qwench
Flour	White Wings	Bright Wings	Baker's Help
Icecream	Drumstick	Icedrum	Delicia
Jam	IXL	XLI	Berie
Milk	Dairy Farmers	Jersey Farmers	Grazers Choice
Muesli Bars	Uncle Tobys	Uncle Homers	Fruity Muse
Mustard	MasterFoods	MasterFeast	FastFlavour
Noodles	Fantastic	Terrific	Nixoes
Pasta	Barilla	Pastilla	Italiano
Peanut Butter	Kraft	Kroft	Smooch
Pet Food	Optimum	Optimal	Agility
Salt	Saxa	Maxa	Seasons
Shampoo	Pantene	Santene	Agent
Soft Drink	Sprite	Brite	Zapfiz
Soup	Continental	European	Slurp 'N Go
Sports Drink	Gatorade	Energade	Actron
Stain Remover	Vanish	Vanquish	Keepoff

Sugar	CSR	CRS	Sweetr
Tea	Bushells	Brambells	Best Brew
Tissues	Kleenex	Cleanix	Clearnose
Yoghurt	Yoplait	Plaityo	Snappo
