Title	Subjective evaluation of running footwear depends on country and assessment method: A bi-national study
Author(s)	Pui W. Kong, Chen Y. Lim, Rui Ding and Thorsten Sterzing
Source	*Ergonomics, 58*(9), 1589-1604. http://dx.doi.org/10.1080/00140139.2015.1018957
Published by	Taylor & Francis (Routledge)

Subjective Evaluation of Running Footwear Depends on Country and Assessment Method: A Bi-national Study

Pui W. Kong[1], Chen Y. Lim[1], Rui Ding[2], Thorsten Sterzing[2]

[1]Physical Education and Sports Science Academic Group, National Institute of Education, Nanyang Technological University, Singapore.

[2]Li Ning Sports Science Research Center, Beijing, China.

Running Head: Running Shoe Perception in China and Singapore

Pui W. Kong, Ph.D.* (Corresponding author)
Physical Education and Sports Science Academic Group, National Institute of Education,
Nanyang Technological University, 1 Nanyang Walk, Singapore, 637616.
Phone: (+65) 6219 6213
Fax: (+65) 6896 9260
E-mail: puiwah.kong@nie.edu.sg

Chen Y. Lim
Physical Education and Sports Science Academic Group, National Institute of Education,
Nanyang Technological University, 1 Nanyang Walk, Singapore, 637616.
E-mail: chenyenlim@hotmail.com

Rui Ding
Li Ning Sports Science Research Center, Li Ning (China) Sports Goods Co., Ltd, No. 8 Xing Guang 5th Street, Beijing, China, 101111.
E-mail: rui.ding@li-ning.com.cn

Thorsten Sterzing, Ph.D.
Li Ning Sports Science Research Center, Li Ning (China) Sports Goods Co., Ltd, No. 8 Xing Guang 5th Street, Beijing, China, 101111.
E-mail: thorsten.sterzing@web.de

Abstract

This study examined 1) the perception of running shoes between China (Beijing) and Singapore, and 2) whether running shoe preference depended on assessment methods. One hundred (n=50 each country) Chinese males subjectively evaluated four shoe models during running using two assessment procedures. Procedure 1 used a Visual Analogue Scale (VAS) to assess five perception variables. Procedure 2 was a ‘Head-to-head’ comparison of two shoes simultaneously (e.g. left foot: A, right foot: B) to decide which model was preferred. VAS scores were consistently higher in Beijing participants (*P*<.001), indicating a higher degree of liking. Singapore participants used the lower end but a wider range of the 15-cm scale for shoe discrimination. Moderate agreement was seen between the VAS and ‘Head-to-head’ procedures, with only 14 out of 100 participants matched all 6 pairwise comparisons (median=4 matches). Footwear companies and researchers should be aware that subjective shoe preference may vary with assessment methods.

Word Count: 150 (abstract)

Key Words: Shoes; Singapore; Beijing; visual analogue scale; head-to-head

Practitioner Summary

Footwear preference depends on country and assessment methods. Running shoe perception differed between Beijing and Singapore Chinese, suggesting that footwear recommendation should be country-specific. Individuals' shoe preference measured by Visual Analogue Scale when wearing complete pairs may not reflect that when directly comparing different models in left and right feet.

Word Count: 50 (practitioner summary)

1. Introduction

Subjective rating is an important component of footwear evaluation, alongside mechanical and biomechanical tests (Hennig 2011). Criteria such as fit, comfort and overall quality of the shoes often provide important information for footwear design (Au and Goonetilleke 2007, Hennig 2011). Much research has been done on footwear comfort with applications in military, sports, clinical and general populations (Burke 2012, Hong et al. 2005, Lake and Lafortune 1998, Llana et al. 2002, Mills, Blanch, and Vicenzino 2012, Mundermann, Stefanyshyn, and Nigg, 2001, Sterzing et al. 2013, Witana, Feng, and Goonetilleke 2004). In everyday life, customers often rely on their subjective 'feeling' and perception of the shoes when comparing various shoe models.

Footwear perception is a complex phenomenon that is influenced by many factors (Lam, Sterzing, and Cheung 2011, Mills, Blanch, and Vicenzino 2010, Mundermann et al. 2001, 2002, Sterzing, Lam, and Cheung, 2012). Shoe preference differs between males and females, and most people prefer the same type of shoes for different activities (Kong and Bagdon 2010). Skeletal alignment such as heel eversion angle measured while standing may affect comfort rating of a shoe during standing, walking and running but the variations among subjects and shoe models are high (Miller et al. 2000). Ratings of perceived cushioning can also be influenced by investigator comments rather than shoe construction characteristics (McCaw, Heil, and Hamill, 2000). A shoe model that is well perceived in one country may not be suitable for another country due to factors such as differences in foot anatomy, climate and cultural background (Bilkey and Nes 1982, Elliott and Cameron 1994, Hawes et al. 1994, Mauch et al. 2008, McBain

1977, Stolwijk et al. 2013). While there are many footwear companies that have retail stores world-wide, the perception of shoes between countries has not been studied.

Measuring footwear perception is difficult. Although some subjective measures are related to biomechanical and/or mechanical properties of the shoes (Lam et al. 2011, Milani, Hennig, and Lafortune 1997, Pisciotta and Shorten 2013, Sterzing et al. 2013), it is still unclear what contributes to the overall footwear perception (Goonetilleke 1999, Llana et al. 2002, Miller et al. 2000, Mills et al. 2010, Mundermann et al. 2003). A recent review based on a series of studies over 18 years concluded that human subjects generally cannot clearly distinguish between various features of a running shoe (e.g. pronation control, shock absorption) (Hennig 2011). If one likes a shoe, he/she tends to like all features of that shoe. In contrast, another study showed that a multiple regression model of four footwear perception factors (fit, cushioning, arch support and stability) explained 77.7% to 81.6% of variance in 'overall preference' in running shoes (Tay et al 2014). In addition, these individual footwear perception factors did not exhibit multicollinearity, suggesting that each factor explained a unique part of the total variance. The general ability to differentiate between shoe stimuli depends on the magnitude of the difference. One study by Pisciotta and Shorten (2013) investigated whether human subjects could correctly rank five pairs of shoes from least cushioned to most cushioned (g-max impact scores 11.1, 12.2, 13.1, 14.6 and 17.6). The threshold of perception was found to be a difference in cushioning of approximately 1 g, with 6% improved perception matching for each additional 1 g difference. This suggests that there exists a certain perception threshold before one can subjectively evaluate differences between shoes.

To assess footwear perception, the Visual Analogue Scale (VAS) is widely used (Clinghan et al. 2008, Lam et al. 2011, Llana et al. 2002, McCaw et al. 2000, Mills, Blanch, and Vicenzino 2011). This method has demonstrated good reliability in assessing comfort during walking, jogging and running, though no sensitivity characteristics of the VAS used are provided (Mills et al. 2010, Mündermann et al. 2002). VAS based perception measurements have also shown to be more stable than Likert scales (Mills et al. 2010). Rating scales, however, are impractical in real life when comparing different shoe models in a retail store. A more realistic scenario would be a direct 'Head-to-head' comparison, deciding over two shoes simultaneously (i.e. one model on each foot). Logically, one would expect that the favoured shoe model determined from 'Head-to-head' comparison should also score higher in VAS. To our knowledge, 'Head-to-head' comparisons have not been studied in previous research and it is unknown whether shoe preference would depend on assessment methods.

The primary purpose of this study was to compare the perception of running shoes between two countries, namely China (Beijing) and Singapore. A secondary aim was to investigate whether running shoe preference depended on assessment methods. It was hypothesized that 1) running shoe perception did not differ between Beijing and Singapore Chinese; and 2) running shoe preference did not differ between VAS and 'Head-to-Head' assessment methods.

2. Methods

2.1 Participants

This study was approved by the Nanyang Technological University Institutional Review Board. Written informed consent was obtained from participants prior to data collection. One hundred (n

= 50 Beijing, n = 50 Singapore) male runners aged between 18 to 35 years with foot sizes within US 8.0 to US 9.5 were recruited in the study (Tables 1 and 2). Each participant completed a survey on demographics, medical history and training routine to determine eligibility. To minimize the influence of ethnicity, the participants and both parents must be Chinese in race. In addition, the participants must be born and lived in their respective areas, either Beijing or Singapore, for at least 15 years, adopt a rearfoot striking pattern, and run at least twice with a total distance of 10 km per week in the past three months. All participants were familiar with treadmill running. Exclusion criteria included current pain in any part of the body, or any injuries at the back and lower extremities within the past six months.

*** Table 1 near here ***

*** Table 2 near here ***

2.2 Instrumentation

Four pairs of Li Ning running shoes of the same size (US size 9.0) were used for treadmill running tests (Table 3). All shoes were black in color to minimise the influence of appearance. One set of new shoes was used for each country, shared by all 50 participants. Initially, the new shoes were broken in by 5 minutes of treadmill running prior to mechanical tests and perception assessments in each country. The mechanical properties of the two sets of shoes were almost identical based on rearfoot impact scores (g).

For the perception test, each participant wore a new pair of running socks during the tests to standardize the foot-shoe interface. The same treadmill model (TechnoGym Excite+ RUN NOW

700) was used in both sites. A standard Brannock device (The Brannock Device Co., Liverpool, NY, USA) was used to measure foot lengths, arch lengths and foot widths (Figure 1). Length measurements were reported in US sizes. Foot widths were coded from 1 (AAA) to 9 (EEE), with a greater value indicating a wider foot (Luximon & Goonetilleke, 2004).

*** Table 3 near here ***

*** Figure 1 near here ***

2.3 Procedures

Identical test procedures were carried out first in Beijing and then in Singapore by the same investigator (CYL) who was bilingual in Mandarin Chinese (for Beijing) and English (for Singapore). Eligible participants warmed up on the treadmill for 4 minutes, running in each pair test shoes for 1 minute. During the warm-up run, participants were asked to adjust the speed to their preferred pace after which the same speed was used throughout the experiment. Two perception assessment procedures were carried out in a randomized order:

Procedure 1: VAS Assessment

Five footwear perception variables were assessed on a horizontal 15-cm VAS scale: ‘Fit’, ‘Cushioning’, ‘Arch Support’, ‘Stability’, and ‘Overall Shoe Preference’ (0: dislike extremely – 15: like extremely). Each variable was assessed on a separate VAS form printed in A4-size with a standard script *“With regard to your shoes, please rate the following by drawing a single vertical line on the scale”*. Samples of VAS forms in both English (Singapore) and Chinese (Beijing) are detailed in the Appendices. The four pairs of test shoes were presented in a mixed order such that all 24 possible sequences (e.g. ABCD, ABDC) were used and shared among the participants. All tests on one shoe model were completed first before commencing the next model. For each shoe model, participants ran at their preferred speed for 5 runs lasting 1 minute each (excluding the treadmill acceleration period). Before each run, participants were informed on which VAS variable they should focus on. After each 1-min run, the treadmill was stopped and the participants were asked to assess the shoe model on the provided VAS variable. The treadmill was then re-started to reach the target speed for the next 1-min run. The assessment orders of the first four variables (‘Fit’, ‘Cushioning’, ‘Arch Support’, ‘Stability’) were selected using a mixed order of presentation to include all possible sequences, while ‘Overall Shoe Preference’ was always administered last. Before the first and in between shoe conditions, participants jogged on the spot on the ground for 10 seconds in socks only to desensitize the feeling of the previous test shoes.

Raw VAS drawings were measured to the nearest 0.1 cm. The minimum, maximum and discrimination range used by each participant were identified. From the ‘Overall Shoe Preference’ measurement, a shoe ranking (e.g. A > B > C > D) was determined. In addition, the relative likings of a shoe versus the other 3 models were identified. There were 6 possible comparisons (i.e. A-B, A-C, A-D, B-C, B-D, C-D) for each participant, resulting in a total of 600 pairwise likings. This method of pairwise comparison could result in a ‘draw’ condition when identical VAS measurements were obtained for two or more shoe models.

Procedure 2: ‘Head-to-head’ Comparison

Procedure 2 consisted of 6 runs lasting 1 minute each, comparing the same four pairs of shoes as in Procedure 1. During each 1-min run, participants wore two different shoe models (e.g. left foot: A, right foot: B). At the end of the run, the treadmill was stopped and the participants were asked to decide which model they preferred over the other (e.g. A over B, or B over A) based on their ‘overall shoe preference‘.When presenting the six mixed shoe pairs (i.e. A-B, A-C, A-D, B-C, B-D, C-D) for *‘head-to-head’* comparison, it was ensured that no two participants were provide with the same pair sequence. For each of the 6 paired comparisons, the left/right foot assignment was mixed to ensure an equal chance for each foot getting a particular shoe. For example, in the comparison of A versus B, there was an equal chance for the left foot to be assigned A or B. Similar to Procedure 1, participants jogged on the spot for 10 seconds in socks only before the first and in between shoe comparisons to desensitize the feeling of the previous test shoes. The number of times a shoe was preferred over another model was counted (maximum 3 per participant) to determine a relative ranking of the 4 models (e.g. A > B > C > D).

2.4 Reliability Tests

To assess the test-retest reliability, a sub-set of participants (n = 13 Beijing, n = 15 Singapore) were invited to repeat exactly the same procedures seven days after the initial test.

2.5 Data Analyses

All analyses were performed using the SPSS software (version 21.0). Statistical significance level was set at $P \leq .05$.

Country

Descriptive statistics are presented as mean (SD) for VAS data. A 4 × 2 mixed design Analysis of Variance (ANOVA) with repeated measures (within-subject factor = shoes, between-subject factor = country) was used to detect differences in each of the five VAS variables. The Mauchly's test was first used to examine whether the assumption of sphericity was violated and if so the Greenhouse-Geisser estimate was used to correct the degrees of freedom. Post-hoc analyses with Bonferroni adjustment were applied when appropriate. The VAS minimum, maximum and discrimination range between Beijing and Singapore were compared using independent t-tests (two-tailed). In order to assess whether the between-country differences were due to different usage and interpretation of the VAS, we performed additional analysis using normalised data. First, the VAS raw scores were normalised into z-scores which were expressed in terms of standard deviations from their means using the following formula:

$$z = \frac{score - mean}{standard\ deviation}$$

Subsequently, we re-ran all statistical analyses for between-country comparisons using the normalised data.

Assessment method

To examine the influence of assessment method on shoe preference, the pairwise likings obtained from the VAS and 'Head-to-head' procedures were compared within each subject. The comparison was considered 'a match' if the relative liking (e.g. A > B) from the VAS 'Overall Shoe Preference' was consistent with that in the 'Head-to-head' comparison. Since the maximum number of matches is 6 for each individual, we considered the agreement as poor for 0 to 3 matches, moderate for 4 matches, good for 5 matches and very good for 6 matches. Combining all subjects together, the total number of matches was counted and expressed as a percentage out of the possible 600 comparisons.

Reliability

The test-retest reliability of the VAS data were assessed using intra-class correlation (ICC) and standard error of measurement (SEM) (Mills, et al., 2010). The ICC coefficients were interpreted as slight (0 - .20), fair (.21 - .40), moderate (.41 - .60), substantial (.61 - .80) and almost perfect (> .80) according to Altman (1991). For the 'Head-to-head' comparison, the number of matches in shoe preference between the two test days was counted (e.g. A > B on both days). Since the maximum number of matches is 6, we considered the reliability as poor for 0 to 3 matches, moderate for 4 matches, good for 5 matches and very good for 6 matches.

3. Results

The preferred running speeds used in the experiments were similar between countries [mean (SD), Beijing =7.40 (1.34) km/h, Singapore = 7.78 (1.33) km/h, P = .163]. For VAS reliability, the ICC ranged from .46 to .90 with the majority greater than .60 (substantial to almost perfect) across all variables and shoes. The SEM ranged from 1.05 to 1.91 cm. For the 'Head-to-head' comparison, the median number of matches in shoe preference between day 1 and day 2 was 4.5 (range = 1 – 6). Collectively, we considered moderate to good reliability for these two methods of footwear perception assessments.

A significant main effect of country (P < .001) and shoes (P < .001) was found in all five VAS variables (Table 4). The raw VAS scores were consistently higher in Beijing than Singapore participants, with between-country significance remaining when data were normalised (all $P <$ .001). Overall, participants preferred model D the most, followed by models B, A and C. One significant shoes × country interaction (P = .033) was found in ‘Stability’, with a much lower VAS value for model C among Singapore participants. Singapore participants used the lower end but a wider range of the 15-cm scale for shoe discrimination in both raw and normalised VAS data analyses (Figure 2). Using the ‘Head-to-head’ method, the overall relative rankings of the shoes were D > B > A > C, with shoe D preferred most in both countries (Table 5).

*** Table 4 near here ***

*** Figure 2 near here ***

*** Table 5 near here ***

The agreement in the pairwise comparisons between the VAS and the ‘Head-to-head’ procedures was considered moderate, with an overall median number of matches of 4.0 (range 0 to 6). Only 14 (Beijing 6, Singapore 8) out of 100 participants matched all 6 pairwise comparisons, 25 (Beijing 13, Singapore 12) matched 5 comparisons, 39 (Beijing 19, Singapore 20) matched 4 comparisons, and 22 matched poorly (3 matches or less) including 1 failed to match any. As a whole group, there were 414 (69%) matches out of the possible 600 comparisons with 5 (0.8%) cases of ‘draws’ resulting from identical VAS scores. The agreement in assessment methods was slightly higher among Singapore (71% match) than Beijing (67% match) participants.

4. Discussion

This study investigated the perception of running shoes between China (Beijing) and Singapore using two assessment procedures. The main findings are: (1) running shoe perception and VAS discrimination range differed between the two countries, and (2) the agreement in shoe rankings derived from the VAS and 'Head-to-head' procedures was moderate.

4.1 Shoe perception differed between Beijing and Singapore

The clear between-country VAS differences were somewhat unexpected. The between-country differences detected even after VAS data normalization reassured that our observations mark true differences and are not influenced by varying interpretation of the scale. Beijing participants consistently rated all shoes higher across all variables than their Singapore counterparts, indicating a higher degree of satisfaction. We speculate a few possible reasons here.

First, the test shoes may fit the foot shape better for the Beijing participants who had wider feet than the Singaporeans (Table 2). We acknowledged that a good fit is critical for comfort (Hennig 2011, Miller et al. 2000) and to support proper function of the features assessed in our research. Despite delimiting our study to only Chinese to minimise ethnic influences (Hawes et al. 1994, Stolwijk et al. 2013), differences in foot anatomy still exist. Due to the hot climate all year round, Singaporeans are usually unshod at home and wear open-toes footwear such as sandals and slippers (flip flops) outdoor. The different environmental factors and footwear habits between countries may have contributed to the different foot shapes (Cheskin 1987, Mauch et al. 2008). It should be noted that despite all participants had similar foot length, Singapore Chinese were shorter and lighter (Table 1), and had narrow feet (Table 2) than Beijing Chinese, the latter explaining differences in fit perception between groups.

Second, we did not blind the brand name in this study because only Li Ning shoes were used. Since Li Ning origins and is a major sports brand in China, the Beijing participants may have a domestic brand preference towards its shoes (Wang 2010). Relatively, Li Ning is not as well established as other sports brands in Singapore, leading to lower perceived ratings of its products (Hennig and Schulz 2013). Third, the difference in social-economic status between the Singapore and Beijing populations may trigger various degrees of product satisfaction among the groups. Singapore participants may be able to afford more high-end running shoes from more renowned global manufacturers. Past experience could have influenced them to be more critical when evaluating a shoe and therefore lead to an overall lower rating of the shoes used in the present study. Although we cannot pinpoint to the exact reason, our findings highlight that running shoe satisfaction depends on country. It is therefore not advisable to generalise footwear perception characteristics determined from one population to another.

It is also interesting to note that the VAS discrimination range was wider in Singapore (mean = 5.9 cm across 5 variables) than Beijing (mean = 4.6 cm). Generally speaking, Singapore is perceived as more liberal and open in culture while China more conservative (Bernstein and Munro 1997, Yeoh and Chang 2001). The more reserved nature of Beijing Chinese may have resulted in a narrower range of shoe discrimination. These possible cultural influences on VAS usage add complexities and difficulties in detecting true differences in the liking of footwear between countries.

Although the VAS scores were higher in Beijing, the pattern of shoe liking sequence (D > B > A > C) is almost entirely the same between the two countries for all five variables (Table 4). This indicates that the relative liking of the shoes is rather similar between countries, even though the degree of satisfaction is higher in Beijing. Similarly, the ‘Head-to-head’ comparisons also showed that shoe model D was the most favoured in both countries, though there are slight discrepancies in the other three models. Collectively, our findings suggest some degree of similarity in the overall preference of the running shoes between China and Singapore. From the manufacturer’s point of view, a popular shoe model in Beijing may also sell well among Singapore Chinese.

4.2 Shoe rankings differed between assessment methods

When assessing the same four models using two different procedures, less than one in seven participants reached the same preference agreement among all shoes. Although the agreement between methods can be considered moderate overall, there were considerable individual variations with poor agreement seen in 22% of the participants. This means that subjective evaluation of running shoe features using VAS may not reflect one’s overall shoe preference when two models are compared directly. The moderate agreement between the perception assessment procedures reinforces previous studies in the literature that testing subjective perception is complex and difficult (Hennig 2011, Kong and Bagdon 2010, Lam et al. 2011, Miller et al. 2000, Mills et al. 2010; Mundermann et al. 2001, 2002, Sterzing et al. 2013). Furthermore, it raises some concerns regarding the common use of VAS in assessing subjective characteristics of footwear in a laboratory setting because the test results do not necessary reflect one’s shoe preference in real life. It is possible that VAS are valid for assessing shoe features and can provide good direction to guide footwear design and development. When asking for an overall decision on shoe preference, however, these very specific functional considerations may only play a minor role for the common runner.

In our ‘Head-to-head’ comparison, participants wore two different shoe models simultaneously, one in each foot. In the assessment of slip resistance of soles, it was easier for human subjects to compare two types of shoe simultaneously when wearing a different shoe on each foot rather than using an evaluation scale (Tisserand, 1985). Previous studies examining shoe-lacing patterns of running shoes and perceived comfort had also used similar procedures by manipulating the tightness of the right shoes while keeping the left shoe as a reference (Hagen and Hennig 2009, Hagen et al.2010). This experimental design has the advantage of minimising the time needed to compare several shoe models and to provide sensation stimuli of both shoes to participants concurrently. Another viable ‘Head-to-head’ procedure that is likely to happen in retail stores would be to wear a complete pair of shoes on both feet, one model after another, and then decide which pair is preferred. This procedure, however, would require longer test duration and would not provide simultaneous sensation stimuli to participants. In summary, footwear companies and researchers should be aware that one’ shoe preference measured by VAS may differ from that when two shoes models were compared directly. Future studies can consider incorporating more realistic consumer experiences such as ‘Head-to-head’ comparisons in footwear perception assessment.

4.3 Shoe preference and mechanical properties

Previous research have shown that runners were able to perceive differences in rearfoot and forefoot hardness (Sterzing et al. 2013) and that cushioning shoes with good shock attenuating properties were generally preferred (Hennig 2011, Kong and Bagdon 2010). When compared with the least favoured model (C), the top ranked model (D) is characterised by heavier mass (+27.0%), increased rearfoot (+13.3%) and forefoot (+31.3%) thickness, higher forefoot stiffness (+21.1%), lower forefoot impact (-4.6%) and higher forefoot energy return (+11.2%) (Table 3). The distinct forefoot characteristics between these two models support the recent findings that a soft forefoot midsole favoured positive perception assessment of cushioning preference of running shoes (Sterzing et al., 2013). Interestingly, the rearfoot impact scores and energy return are similar between these two extreme shoe models. Considering rearfoot cushioning has been receiving much attention in the past decades, perhaps more future work on forefoot characteristics can help clarify the complex relationships between mechanical properties and subjective perception of athletic footwear (Hennig 2011).

4.4 Limitations

There are some limitations of our study. First, the Beijing participants were taller, heavier, ran more frequently and more mileage than the Singapore participants despite them meeting our inclusion criteria on demographics and running mileage. Second, there may be selection bias in subject recruitment. The Singapore participants were mostly general university students compared to the Beijing Sports University students in China who may have more knowledge and experience in choosing athletic shoes. Third, the experiments were conducted during winter in Beijing whereas Singapore is hot and humid throughout the year. The different in climate may have played a role in influencing foot sensation. Schlee and colleagues (2009) demonstrated that small temperature changes (5 to 6ºC) in the foot sole can influence plantar foot vibration sensitivity, with a lower threshold for warmer skin. Since the treadmill runs took place in an indoor heated laboratory in Beijing and an air-conditioned gymnasium in Singapore respectively, the influence due to foot climate, while likely to exist, is not regarded substantial. Finally, the study was conducted in different languages between China and Singapore. We have minimised the potential influence of language by having the same bilingual investigator to conduct the study in both locations.

5. Conclusions

Running shoe perception differed between Beijing and Singapore Chinese regarding the degree of satisfaction and in part the relative rankings. Moderate agreement in shoe preference was seen when evaluating the same shoes using different assessment procedures. Footwear companies and researchers should be aware that one' shoe preference measured by VAS may not reflect their overall preference decision when two models were compared directly. Future studies can consider incorporating more realistic consumer experiences such as 'Head-to-head' comparisons in footwear perception assessment.

Word Count: 4213 (main text)

Acknowledgements: This work was supported by the Li Ning Sports Science Research Center, Beijing, China. The authors would like to thank Dr Masato Kawabata for his helpful discussion on data analysis.

Disclosure statement: Rui Ding and Thorsten Sterzing were employees of Li Ning (China) Sports Goods Co., Ltd during the time of this research. Pui Kong and Chen Yen Lim declared that there is no conflict of interest.

References

Altman, D.G. 1991. "Practical statistics for medical research." Padstow, Cornwall: Chapman & Hall.

Au, E. Y., and Goonetilleke, R. S. 2007. "A qualitative study on the comfort and fit of ladies' dress shoes." *Applied Ergonomics* 38(6), 687-696.

Bernstein, R., and Munro, R. H. 1997. "The coming conflict with America." *Foreign Affairs* 76(2), 18-35.

Bilkey, W. J., and Nes, W. 1982. "Country-of-origin effects on product evaluations." *Journal of International Business Studies* 13(1), 89-99.

Burke, J. R. 2012. "Effects of footwear comfort perception on the neuromuscular control of balance." *The International Journal of Neuroscience* 122(4), 209-220.

Cheskin, M. P. 1987. "Foot types". In *The complete handbook of athletic footwear,* edited by M. P. Chekin, K. J. Sherkin and B. T. Bates, 234-236. New York, NY: Fairchild Publications.

Classbase. 2012. "Educational Database." Accessed 2013 April 17. http://www.classbase.com/countries

Clinghan, R., Arnold, G. P., Drew, T. S., Cochrane, L. A., and Abboud, R. J. 2008. "Do you get value for money when you buy an expensive pair of running shoes?" *British Journal of Sports Medicine* 42(3), 189-193.

Elliott, G. R., and Cameron, R. C. 1994. "Consumer perception of product quality and the country-of-origin effect." *Journal of International Marketing* 2(2), 49-62.

Goonetilleke, R. S. 1999. “Footwear cushioning: relating objective and subjective measurements.” *Human Factors* 41(2), 241-256.

Hagen, M., and Hennig, E.M. 2009. “Effects of different shoe-lacing patterns on the biomechanics of running shoes.” *Journal of Sports Sciences* 27(3), 267-275.

Hagen, M., Homme, A., Umlauf, T., and Hennig, E.M. 2010. “Effects of difference shoe-lacing patterns on dorsal pressure distribution during running and perceived comfort.” *Research in Sports Medicine* 18, 176-187.

Hawes, M. R., Sovak, D., Miyashita, M., Kang, S., Yoshihuku, Y., and Tanaka, S. 1994. “Ethnic differences in forefoot shape and the determination of shoe comfort.” *Ergonomics* 37(1), 187-196.

Hennig, E. M. 2011. “Eighteen years of running shoe testing in Germeny - a series of biomechanical studies.” *Footwear Science* 3(2), 71-81.

Hennig, E. M., and Schulz, J. 2013. “Subjective evaluation of biomechanical shoe properties during blinded and non-blinded running”. *Footwear Science* 5(Suppl1), S75.

Hong, W. H., Lee, Y. H., Chen, H. C., Pei, Y. C., and Wu, C. Y. 2005. “Influence of heel height and shoe insert on comfort perception and biomechanical performance of young female adults during walking.” *Foot & Ankle International* 26(12), 1042-1048.

Kong, P. W., and Bagdon, M. 2010. “Shoe preference based on subjective comfort for walking and running.” *Journal of the American Podiatric Medical Association* 100(6), 456-462.

Lake, M. J., and Lafortune, M. A. 1998. “Mechanical inputs related to perception of lower extremity impact loading severity.” *Medicine and Science in Sports and Exercise* 30(1), 136-143.

Lam, W. K., Sterzing, T., and Cheung, J. T. 2011. “Reliability of a basketball specific testing protocol for footwear fit and comfort perception.” *Footwear Science* 3(3), 151-158.

Llana, S., Brizuela, G., Dura, J. V., and Garcia, A. C. 2002. “A study of the discomfort associated with tennis shoes.” *Journal of Sports Sciences* 20(9), 671-679.

Luximon, A., and Goonetilleke, R. S. 2004. “Foot shape modeling.” *Human Factors* 46(2), 304-315.

McCaw, S. T., Heil, M. E., and Hamill, J. 2000. “The effect of comments about shoe construction on impact forces during walking.” *Medicine and Science in Sports and Exercise* 37(7), 125-164.

Mauch, M., Mickle, K. J., Munro, B. J., Dowling, A. M., Grau, S., and Steele, J. R. 2008. “Do the feet of German and Australian children differ in structure? Implications for children's shoe design.” *Ergonomics* 51(4), 527-539.

McBain, N. S. 1977. “Developing country product choice: footwear in Ethiopia.” *World Development* 5, 829-838.

Milani, T. L., Hennig, E. M., and Lafortune, M. A. 1997. “Perceptual and biomechanical variables for running in identical shoe constructions with varying midsole hardness.” *Clinical Biomechanics* 12(5), 294-300.

Miller, J. E., Nigg, B. M., Liu, W., Stefanyshyn, D. J., and Nurse, M. A. 2000. Influence of foot, leg and shoe characteristics on subjective comfort. *Foot & Ankle International* 21(9), 759-767.

Mills, K., Blanch, P., and Vicenzino, B. 2010. “Identifying clinically meaningful tools for measuring comfort perception of footwear.” *Medicine and Science in Sports and Exercise,* 42(10), 1966-1971.

Mills, K., Blanch, P., and Vicenzino, B. 2011. “Influence of contouring and hardness of foot orthoses on ratings of perceived comfort.” *Medicine and Science in Sports and Exercise,* 43(8), 1507-1512.

Mills, K., Blanch, P., and Vicenzino, B. 2012. “Comfort and midfoot mobility rather than orthosis hardness or contouring influence their immediate effects on lower limb function in patients with anterior knee pain.” *Clinical Biomechanics* 27(2), 202-208.

Mundermann, A., Nigg, B. M., Humble, R. N., and Stefanyshyn, D. J. 2003. “Orthotic comfort is related to kinematics, kinetics, and EMG in recreational runners.” *Medicine and Science in Sports and Exercise,* 35(10), 1710-1719.

Mundermann, A., Nigg, B. M., Stefanyshyn, D. J., and Humble, R. N. 2002. “Development of a reliable method to assess footwear comfort during running.” *Gait & Posture* 16(1), 38-45.

Mundermann, A., Stefanyshyn, D. J., and Nigg, B. M. 2001. “Relationship between footwear comfort of shoe inserts and anthropometric and sensory factors.” *Medicine and Science in Sports and Exercise* 33(11), 1939-1945.

Pisciotta, E. J., and Shorten, M. 2013. “Perception of running shoe cushioning: wear test vs controlled experiment.” *Footwear Science* 5(Suppl1), S136.

Schlee, G., Sterzing, T., and Milani, T. L. 2009. “Foot sole skin temperature affects plantar foor sensitibity”. *Clinical Neurophysiology* 120(8), 1548-1551.

Sterzing, T., Lam, W. K., and Cheung, J. T.M. 2012. “Athletic footwear research by inductry and academic”. In *The Science of Footwear,* edited by R. S. Goonetilleke, 605-622. CRC Press Taylor & Francis Group.

Sterzing, T., Schweiger, V., Ding, R., Cheung, J. T., and Brauner, T. 2013. "Influence of rearfoot and forefoot midsole hardness on biomechanical and perception variables during heel-toe running." *Footwear Science* 5(2), 71-79.

Stolwijk, N. M., Duysens, J., Louwerens, J. W., van de Ven, Y. H., and Keijsers, N. L. 2013. "Flat feet, happy feet? Comparison of the dynamic plantar pressure distribution and static medial foot geometry between Malawian and Dutch adults." *PloS One* 8(2), e57209.

Tay, S.T., Sterzing, T., Lim, C.Y., Ding, R., and Kong, P.W. 2014. "Multiple regression of running footwear perception factors on overall preference." *Chinese Journal of Sports Biomechanics* 10(S1), 9.

Tisserand, M. "Progress in the prevention of falls caused by slipping." *Ergonomics 28(*7), 1027-1042.

Wang, J. 2010. *Brand new China: advertising, media and commercial culture*. USA: Harvard University Press.

Witana, C. P., Feng, J., and Goonetilleke, R. S. 2004. "Dimensional differences for evaluating the quality of footwear fit." *Ergonomics* 47(12), 1301-1317.

Yeoh, B. S. A., and Chang, T. C. 2001. "Globalising Singapore: debating transnational flows in the city." *Urban Studies* 38(7), 1025-1044.

Figure and Table Captions

Figure 1. Brannock foot device indicating right foot measurements for the arch length, foot size and foot width.

Figure 2. Comparisons of Visual Analogue Scale (VAS) scores and discrimination ranges (cm) between Beijing and Singapore Chinese by independent t-tests. *Raw VAS scores significantly differed from Beijing ($P < .05$). #Normalised VAS scores significantly differed from Beijing ($P < .05$).

Table 1. Physical characteristics and running routine of the participants

Table 2. Brannock foot measurements of the participants

Table 3. Mechanical characteristics of the four Li Ning test shoes

Table 4. Visual Analogue Scale (VAS) measurements (cm) for 4 shoe models

Table 5. Number of times (percentage) a shoe was preferred in the 'Head-to-head' comparisons out of 150 comparisons per shoe each per country and 300 combined

Table 1. Physical characteristics and running routine of the participants

	Beijing	Singapore
n	50	50
Age (yrs)	23.1 (2.2)	23.9 (2.9)
Height (cm)	175.4 (4.3)	171.4 (5.2)*
Body mass (kg)	70.2 (7.1)	64.4 (7.4)*
Running frequency per week	4.2 (1.3)	2.5 (1.4)*
Runner mileage (no. of participants)		
Less than 5 km per week	0	20
5 to 10 km per week	22	20
More than 10 km per week	25	7
Not sure	3	3

*Significantly differed from Beijing ($P < .05$) detected by independent t-test.

Table 2. Brannock foot measurements of the participants

		Beijing	Singapore
Foot length (US size)	Left	8.6 (0.5)	8.6 (0.5)
	Right	8.4 (0.4)	8.6 (0.5)
Arch length (US size)	Left	9.3 (0.8)	9.2 (0.8)
	Right	9.4 (0.5)	9.3 (0.7)
Foot width*	Left	6.2 (0.9)	5.9 (0.9)
	Right	6.4 (0.8)	6.0 (0.9)

*Significant difference between left and right feet ($P = .001$) and between Beijing and Singapore ($P < .001$) detected by repeated measures ANOVA (leg × country).

Table 3. Mechanical characteristics of the four Li Ning test shoes

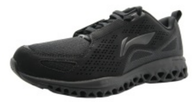
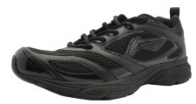
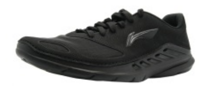
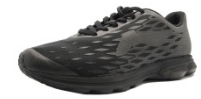
Characteristics	Running shoe model			
	Hyper Arc (A)	Basic Cushion (B)	Superlight (C)	Unit Bow (D)
				
Length (mm)	2750	2750	2750	2750
Mass (kg)	0.327	0.295	0.230	0.292
Forefoot Thickness (mm)	19.0	20.0	16.0	21.0
Rearfoot Thickness (mm)	32.0	31.0	29.0	32.0
Rearfoot to Forefoot Pitch (mm)	13.0	11.0	13.0	11.0
Forefoot Midsole Width (mm)	11.1	11.5	10.6	10.7
Rearfoot Midsole Width (mm)	8.7	8.9	8.5	8.4
Forefoot Stiffness (Nm/deg)	0.19	0.16	0.19	0.23
Forefoot Impact Score (g)	16.7	16.3	15.2	14.5
Rearfoot Impact Score (g)	11.6	9.8	11.1	11.2
Forefoot Energy return (%)	50.3	57.3	50.1	55.7
Rearfoot Energy Return (%)	44.4	46.0	45.7	46.6

Table 4. Visual Analogue Scale (VAS) measurements (cm) for 4 shoe models

VAS variable	Country	Shoe A	Shoe B	Shoe C	Shoe D
Fit	Beijing	9.7 (2.9)	9.9 (2.9)	8.9 (3.4)	10.9 (3.0)
	Singapore	8.4 (3.0)	8.0 (3.0)	7.5 (3.1)	9.8 (2.4)
	Combined	9.0 (3.0)[a]	9.0 (3.1)[ab]	8.2 (3.3)[ab]	10.3 (2.7)
Cushioning	Beijing	9.1 (3.1)	11.0 (2.9)	8.2 (3.4)	10.5 (3.0)
	Singapore	8.3 (2.9)	9.6 (2.4)	5.8 (3.5)	8.7 (2.9)
	Combined	8.7 (3.0)	10.3 (2.7)[a]	7.0 (3.6)	9.6 (3.1)[a]
Arch Support	Beijing	9.4 (2.8)	10.0 (3.0)	8.4 (3.3)	10.5 (2.5)
	Singapore	8.3 (2.9)	7.9 (3.2)	5.9 (3.1)	8.6 (3.0)
	Combined	8.9 (2.9)[a]	8.9 (3.3)[ab]	7.1 (3.4)	9.5 (2.9)[ab]
Stability	Beijing	10.0 (3.1)	10.0 (2.8)	9.7 (2.7)	11.1 (2.2)
	Singapore	8.5 (2.7)	9.2 (2.4)	7.0 (3.2)	9.0 (2.7)
	Combined	9.2 (3.0)[a]	9.6 (2.7)[ab]	8.4 (3.2)[a]	10.0 (2.6)[ab]
Overall Shoe Preference	Beijing	9.7 (3.1)	11.2 (2.7)	8.8 (3.4)	11.3 (2.7)
	Singapore	8.6 (2.8)	9.0 (2.8)	7.0 (3.5)	9.6 (2.9)
	Combined	9.2 (3.0)	10.1 (3.0)[a]	7.9 (3.6)	10.4 (2.9)[a]

Note: *post-hoc* results are indicated in the 'combined' data. Superscripts denote statistically homogenous group. For example, any condition with an "a" superscript is not statistically different from other conditions with an "a" but differed from those without an "a".
A significant main effect of country ($P < .001$) and shoes ($P < .001$) was found in all 5 VAS variables. Between-country difference remained when normalised VAS data were used (all $P < .001$)

Table 5. Number of times (percentage) a shoe was preferred in the 'Head-to-head' comparisons out of 150 comparisons per shoe each per country and 300 combined

Country	Shoe A	Shoe B	Shoe C	Shoe D
Beijing	51 (34%)	90 (60%)	52 (35%)	107 (71%)
Singapore	80 (53%)	78 (52%)	52 (35%)	90 (60%)
Combined	131 (44%)	168 (56%)	104 (35%)	197 (66%)

Note: A total of 600 comparisons were made (300 per country), of which each shoe could be preferred for up to 300 times (150 per country).

Appendices

Appendix A: English Visual Analogue Scale on Fit

Appendix B: English Visual Analogue Scale on Cushioning

Appendix C: English Visual Analogue Scale on Arch Support

Appendix D: English Visual Analogue Scale on Stability

Appendix E: English Visual Analogue Scale on Overall Shoe Preference

Appendix F: Chinese Visual Analogue Scale on Fit

Appendix G: Chinese Visual Analogue Scale on Cushioning

Appendix H: Chinese Visual Analogue Scale on Arch Support

Appendix I: Chinese Visual Analogue Scale on Stability

Appendix J: Chinese Visual Analogue Scale on Overall Shoe Preference

Appendix A: English Visual Analogue Scale on Fit

With regard to your shoes, please rate the following by drawing a single vertical line on the scale:

FIT

Dislike Extremely ______________________________ Like Extremely

Appendix B: English Visual Analogue Scale on Cushioning

With regard to your shoes, please rate the following by drawing a single vertical line on the scale:

CUSHIONING

Dislike Extremely ________________________________ Like Extremely

Appendix C: English Visual Analogue Scale on Arch Support

With regard to your shoes, please rate the following by drawing a single vertical line on the scale:

ARCH SUPPORT

Dislike Extremely	________________________________	Like Extremely

Appendix D: English Visual Analogue Scale on Stability

With regard to your shoes, please rate the following by drawing a single vertical line on the scale:

STABILITY

Dislike Extremely	________________________________	Like Extremely

Appendix E: English Visual Analogue Scale on Overall Shoe Preference

With regard to your shoes, please rate the following by drawing a single vertical line on the scale:

OVERALL SHOE PREFERENCE

Dislike Extremely ____________________________________ Like Extremely

Appendix F: Chinese Visual Analogue Scale on Fit

对您现在穿的鞋子，请在量表上划一条竖线来评价下面的指标

合脚性

非常 不喜欢	______________________________	非常 喜欢

Appendix G: Chinese Visual Analogue Scale on Cushioning

对您现在穿的鞋子，请在量表上划一条竖线来评价下面的指标

缓冲性

非常 不喜欢	______________________	非常 喜欢

Appendix H: Chinese Visual Analogue Scale on Arch Support

对您现在穿的鞋子，请在量表上划一条竖线来评价下面的指标

足弓支撑

非常 不喜欢	________________________________	非常 喜欢

Appendix I: Chinese Visual Analogue Scale on Stability

对您现在穿的鞋子，请在量表上划一条竖线来评价下面的指标

稳定性

非常 不喜欢	______________________________	非常 喜欢

Appendix J: Chinese Visual Analogue Scale on Overall Shoe Preference

对您现在穿的鞋子，请在量表上划一条竖线来评价下面的指标

总体评价

非常 不喜欢	______________________________	非常 喜欢