



Research Article

IJSEHR 2020; 4(1): 1-6
© 2020, All rights reserved
www.sportscienceresearch.com
Received: 09-02-2020
Accepted: 03-04-2020

Application of t-distributed Stochastic Neighbor Embedding (t-SNE) to clustering of social affiliation and recognition psychological motivations in masters athletes

Joe Walsh¹, Ian Timothy Heazlewood², Mark DeBeliso³, Mike Climstein^{4,5}

¹ Sport Science Institute www.sportscienceinstitute.com

² The School of Psychological and Clinical Sciences, Faculty of Engineering, Health, Science and the Environment, Charles Darwin University, Darwin, Northern Territory, Australia

³ Department of Physical Education and Human Performance, Southern Utah University, Cedar City, USA

⁴ Clinical Exercise Physiology, Southern Cross University, School of Health and Human Sciences, Gold Coast, Queensland, Australia

⁵ Physical Activity, Lifestyle, Ageing and Wellbeing Faculty Research Group, University of Sydney, Sydney, NSW, Australia, 2006

Abstract

An exploration of clustering of psychological motivations for participation in sport was conducted using t-distributed Stochastic Neighbor Embedding (t-SNE). The data source used for this investigation was survey data gathered on World Masters Games competitors using the Motivations of Marathoners Scales (MOMS). The aim of this research was to assess the suitability of applying t-SNE to creating two-dimensional scatter plots to visualise the relationship between different psychological motivators for the Social Motives category of the MOMS. Application of t-SNE plots could assist in visually mapping psychological constructs and gaining greater understanding of the underlying patterns in the MOMS tool. Although there was more disparity in the clustering of categories within Social Motives than was hypothesised, some clustering patterns were observed. Some items in the MOMS Social Motives category were connected in a logical manner that complied with those originally proposed by the developers of the MOMS. Two-dimensional scatter plots produced using t-SNE may assist in creating hypotheses about the relationships present between psychological constructs in such high-dimensional data.

Keywords: Social Motives, Affiliation, Recognition, Motivations of Marathoners Scales, Masters Athletes.

INTRODUCTION

The World Masters Games

This manuscript focuses on exploration of clustering of scores from psychometric data gathered on masters athletes. Masters athletes are defined as those systematically training for and competing in organized sporting events designed specifically for older adults [1]. Competing at sport in older ages has been shown to be beneficial for a number of health indices including general cardiovascular health [2], blood pressure [3], improved lipids [4], reduced frailty/sarcopenia [5] and muscular strength and function [6]. The biggest masters sporting event (by participant number) is the World Masters Games (WMG). Participation at the WMG is open to sports people of all abilities, limited by age. The minimum age criterion ranges between 25 and 35 years depending upon the sport. The data used in this manuscript was data gathered at the Sydney WMG, which attracted 28,089 competitors who represented 95 countries competing in 28 sports [7-9]. Research on the masters athletes competing at the Sydney WMG has included investigation of smoking prevalence [10], body mass index [7, 11-18], injury incidence [19-23] and health [4, 24-31] of competitors.

The Motivations of Marathoners Scales

The Motivations of Marathoners Scales (MOMS) (37) is a psychometric instrument based upon a series of 56 questions and scored on a seven-point Likert scale (35). To complete the MOMS participants rate each of the following items according to the scale in terms of how important it was as a reason for their

*Corresponding author:

Dr. Mike Climstein

Clinical Exercise Physiology,
B7.39, Southern Cross
University (Gold Coast
campus), School of Health and
Human Sciences, Bilinga,
Queensland, Australia 4225

Email:

michael.climstein@scu.edu.au

Tel: +617-55893330

participation in sport. A score of 1 would indicate that the item is "not a reason" for participation, whereas a score of 7 indicates that the item is a "very important reason" for participation and scores in-between these extremes represented relative degrees of each reason. The following are sample questions which sought responses to word stems such as; "to control my weight", "to compete with others", "to earn respect of peers", "to improve my sporting performance", "to earn respect of people in general", "to socialize with other participants", "to improve my health", "to compete with myself", "to become less anxious", "to improve my self-esteem" and "to become less depressed." A full list of the 56 questions in the MOMS scale and summary statistics for the MOMS scale data gathered at the Sydney WMG has been previously published [9, 32].

The MOMS is a valid and reliable, quantitative instrument for gauging the importance of a range of psychological factors in determining motivations for sports participation. Participant motivation evaluates those factors that enhance or inhibit motivation to participate and are represented by factors such as health orientation, weight concern/weight loss and personal goal achievement [32-34]. The questions in the scale are split into general categories and these are further subset into Scales [32]. For example, for questions in the category Social Motives, 'To socialize with other participants', 'To meet people', 'To visit with friends' and 'To share a group identity with other participants', comprise the Affiliation subset of the Social Motives questions. The other subset of Social Motives questions, Recognition is composed of questions 'To earn respect of peers', 'People look up to me', 'Brings me recognition' and 'To make my family or friends proud of me' [32].

The MOMS scale has been adopted to investigate athletes competing in other sports (other than marathon), including at both multi-sport events [35, 36] and individual sports tournaments such as rugby [37], or triathlon [38] (with some adaption). Data collected using the MOMS scale has also been used as a convenience sample for demonstrating applications of data mining techniques that can be used in exercise science and exercise psychology [39-42].

The age ranges in the research used to develop the MOMS survey instrument had significant overlap with age ranges of participants at the WMG. The questions identified in the MOMS have been demonstrated [43-46] as important motivational constructs and have been used by sport psychology researchers for more than 25 years. A number of studies have been conducted on the MOMS in the context of masters athletes [11, 15, 35, 47-55]. Heazlewood and colleagues [56] re-evaluated the first and second order factor structure of the MOMS instrument with masters athletes, the factor structure identified in the original MOMS instrument was not reproduced with the WMG male and female cohorts.

t-distributed Stochastic Neighbor Embedding

There are a number of established techniques for visualizing high dimensional data. A relatively modern technique that has a number of advantages over many earlier approaches is t-distributed Stochastic Neighbor Embedding (t-SNE) [57]. With t-SNE high dimensional data can be converted into a two-dimensional scatter plot via a matrix of pair-wise similarities.

Stochastic Neighbor Embedding (SNE) converts Euclidean distances between data points into conditional probabilities that represent similarities [58]. In t-SNE the SNE cost function is replaced with a symmetrized version with simpler gradients [57] and t-SNE uses a Cauchy Distribution (one dimensional Student's-t distribution (as opposed to a Gaussian distribution)) to compute the similarity between two points in the lower-dimensional space [57]. This distribution allows for more dispersion in the lower-dimensional space. Similar to SNE, the t-SNE algorithm develops a probability distribution between

factor pairs in the higher-dimensional space with higher probabilities assigned to pairs with higher similarity. A similar probability distribution is then developed in a lower-dimensional map and the Kullback-Leibler divergence [59] between the two distributions is then minimized with respect to the points in the maps using gradient descent. The aim is developing a lower dimensional mapping (in our case two dimensions) where this mapping retains the similarities that were present in the higher dimensional data. The cost function for t-SNE is not convex, thus initializing scripts with different random seed values will result in differing outcomes.

AIM

Effective visualization of data plays a crucial role in knowledge discovery [60]. The MOMS scale contains complex, multi-dimensional relations between 56 different questions, split into a factor structure that has not been replicated in previous research on WMG athletes [56]. The aim of this research was to assess the suitability of applying t-SNE to creating two-dimensional scatter plots to visualise the relationship between different categories of Social Motives, namely Affiliation and Recognition in the MOMS. If suitable plots could be constructed these could assist in visually mapping psychological constructs and gaining greater understanding of the underlying patterns in the MOMS scale. Two-dimensional scatter plots produced using t-SNE may assist in creating hypotheses about the relationships present between Social Motives as psychological constructs in such high-dimensional data. It was hypothesised that with such a large sample from the WMG athletes, there would be visible clustering in the t-SNE graphs based upon the two Social Motives categories Affiliation and Recognition.

METHODOLOGY

Data was collected on masters athletes participating in the Sydney WMG, after approval for the project was granted by a university Research Ethics Committee in accordance with the ethical standards of the Helsinki Declaration of 1975 (revised in 2008) and the Sydney World Masters Games Organising Committee. An online survey was created using Limesurvey, an open-source, web-based application to deliver the survey. The survey consisted of several sections, one of which was the MOMS survey. A total of 3,928 masters athletes completed all 56 questions in the MOMS. This manuscript analyses psychological participation factors contained within the survey. Further details about the survey methodology and an overview of findings from the survey has been previously published [61].

The psychological participation factors included in the survey were 56 questions based on the MOMS [9]. These were analysed using the t-SNE package included in the scikit-learn python machine learning library [62]. Analysis was conducted using Python 3.6.5 using operating system x86_64-apple-darwin15.6.0 (64-bit). After provisional exploratory analysis of different hyperparameters, it was deemed appropriate to keep the majority of t-SNE hyperparameters fixed at their default settings (the standard settings within the scikit-learn library, with default values and a description of each hyperparameter reported in Table 1) and tune the learning rate hyperparameter. The learning rate was tuned from values of 0.0001 to 5000, which was outside the recommended range in the scikit-learn [62] package recommendations (10-1000) [63]. The fixed values for the other main hyperparameters for t-SNE implemented via scikit-learn [63] are listed in Table 1 below.

Table 1: Descriptions and default values for the t-SNE hyperparameters in the scikit-learn package [63]

Hyperparameter	Description	Default value
Number of components	The number of components is the dimension of the embedded space, in our case we generate a two-dimensional space, so we keep the default value.	2
Perplexity	Perplexity is related to the number of nearest neighbours used in other learning algorithms such as k-nearest neighbours [64]	30
Early exaggeration	Early exaggeration is related to the space between clusters in the embedded space where there was already clustering in the original space	12
Number of iterations	This is the maximum number of iterations for the optimisation	1000
Number of iterations without progress	This is the maximum number of iterations without progress before the optimization is aborted	300
Minimum gradient norm	If the gradient norm is below this value the optimization will be halted	1×10^{-7}
Metric	This is the metric to use when calculating distances in a feature array. In our scenario we use Euclidean distance.	The default metric of Euclidean distance was used
Initialization of embedding	Whether to use a random initialization, principle components analysis or an array to initialize embedding	Random initialization was used
Method	This is the gradient calculation method	The default method, Barnes-Hut t-SNE [57] was used
Angle	Angle is a speed versus accuracy trade off hyperparameter	The default value of 0.5 was retained as the Barnes-Hut t-SNE is not very sensitive to changes in this metric [63]

RESULTS

Figures 1-5 display t-SNE graphs produced for a selection of learning rates from 10 to 5000. Additional learning rates were utilised within the range 0.0001 to 5000, however for concise reporting only a selection of these are reproduced in this manuscript.

The t-SNE graph produced with a learning rate of 10 is displayed in Figure 1. Three of the four Social Motives categorised under Affiliation are clustered together in the fourth quadrant of the graph (lower right corner). However, the question "To visit with friends" is separated from these other Affiliation category questions. The Affiliation category questions that were closest in proximity were "To socialize with other participants" and "To meet people".

The t-SNE graph produced with a learning rate of 100 is displayed in Figure 2. Two of the four Social Motives categorised under Affiliation are clustered together in the first quadrant of the graph (upper right corner). These two questions have the highest values for t-SNE dimension 1, whilst the other two Affiliation questions have the lowest values for t-SNE dimension 1. The Recognition questions are clustered closer to the zero value for t-SNE dimension 1, although there is no apparent differentiation between Affiliation and Recognition category questions in terms of t-SNE dimension 2.

The Affiliation questions that were grouped together in pairs were "To socialize with other participants" and "To meet people" as one pair of questions. The second pair were the questions "To visit with friends" and "To share a group identity with other people".

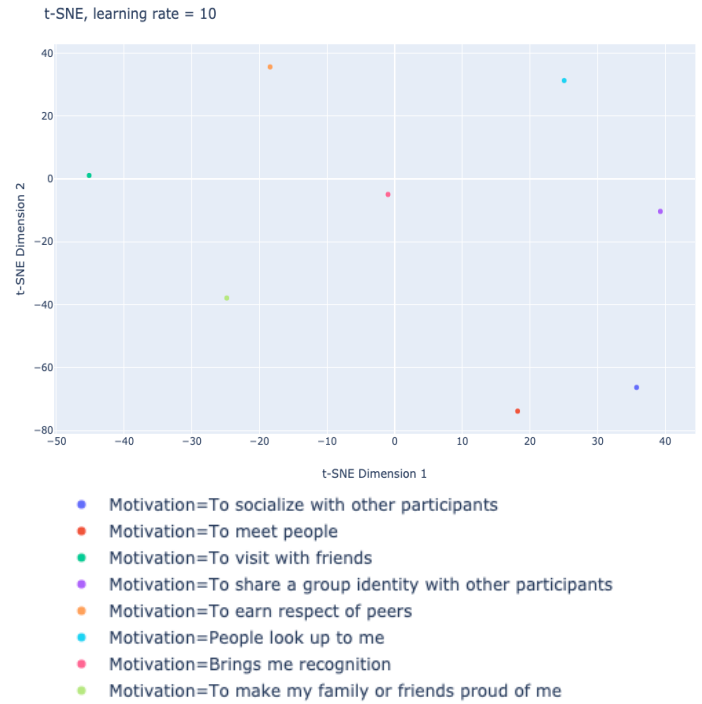


Figure 1: t-SNE graph with Learning Rate 10

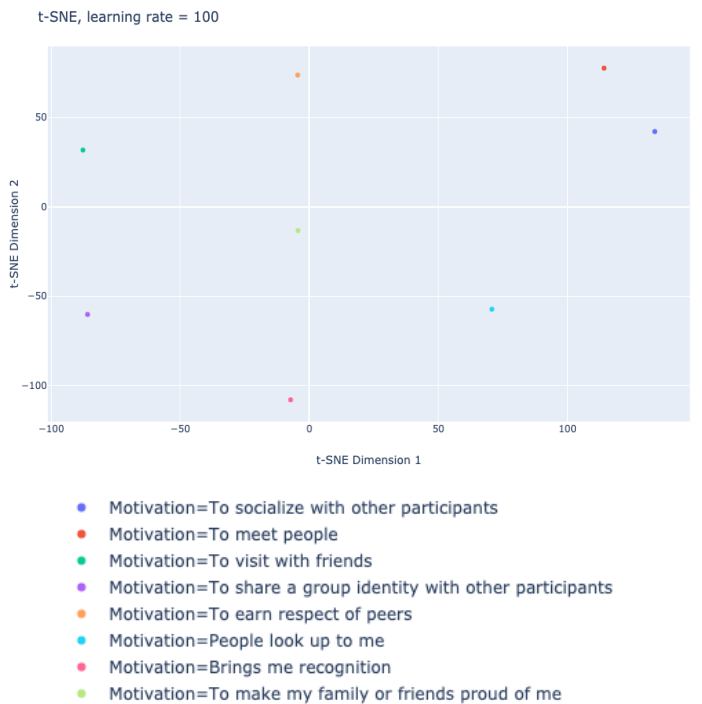


Figure 2: t-SNE graph with Learning Rate 100

The t-SNE graph produced with a learning rate of 500 is displayed in Figure 3. Similar to the previous graph, the Affiliation questions are clustered into two pairs of questions. The questions "To socialize with other participants" and "To meet people" were grouped together. The second pair were the questions "To visit with friends" and "To share a group identity with other people". These pairings are consistent with the pairings demonstrated in the previous graphs.

The t-SNE graph produced with a learning rate of 1000 is displayed in Figure 4. The four Affiliation category questions are located in or near

to the second quadrant of the graph (upper left portion). The pair of questions "To socialize with other participants" and "To meet people" were both located in the second quadrant, whilst the other two Affiliation category questions "To visit with friends" and "To share a group identity with other people" were positioned closer to the centre of the graph, outside this quadrant.

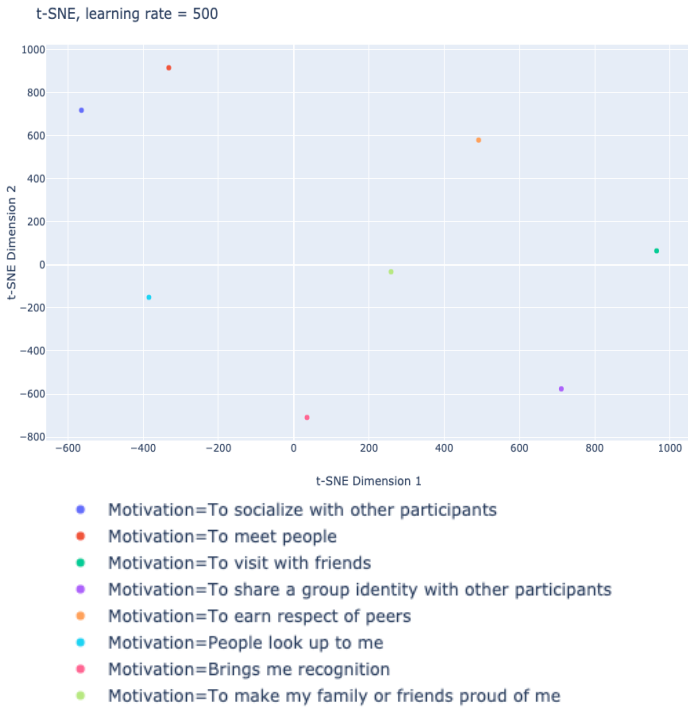


Figure 3: t-SNE graph with Learning Rate 500



Figure 4: t-SNE graph with Learning Rate 1000

The t-SNE graph produced with a learning rate of 2000 is displayed in Figure 5. There was some clustering of questions for the Affiliation category in terms of t-SNE dimension 2. Three of the affiliation category question had positive values for t-SNE dimension 2. The Recognition category questions had lower values for t-SNE dimension 2 than these three Affiliation category questions. The Affiliation category question "To share a group identity with other people" had the most negative value in terms of t-SNE dimension 2 and was isolated

from the other three questions in this category in terms of t-SNE dimension 2.



Figure 5: t-SNE graph with Learning Rate 2000

DISCUSSION

The Figures 1-5 are a visual representation of the clustering of eight of the 56 psychological motivations documented in the literature [9, 32, 46, 65]. As the dimensional reduction utilized in t-SNE is non-linear, the axes in the graphs in Figures 1-5 represent distances in the two-dimensional space, however relating these to equivalent distances in the initial 8 dimensions is a non-linear transformation. Thus, the figures should be used as a visualization tool, however the interpretability in the units of the initial eight dimensional data is not apparent or suitable from the figures. The Cartesian coordinates of different questions was not the focus of this manuscript as t-SNE was utilized to explore the data in terms of Euclidean distance between the questions (as described in the introduction section).

In terms of visualization of relationship between the eight variables there were clearly patterns of clustering which may give insight into relationships within the data. This manuscript focuses upon the Social Motives questions in the MOMS. These questions were utilised as an example of the replication (or disparity) of clustering relationships in the original development of the MOMS instrument [32] when questions are inspected graphically utilising t-SNE.

Inspection of clustering of questions on the t-SNE scatter plots revealed some patterns that were representative of underlying relationships between the different questions. Few of the clustering relationships as proposed in the original scale [32], splitting the Social Motives questions into Affiliation and Recognition were dramatically evident in this data explored using t-SNE. The lack of distinct and consistent clustering in both t-SNE dimensions based on these two categories, does not necessarily mean that these categories are not valid. The disparity in results observed compared to hypothesised trends should be reported, but this is manuscript is an exploratory analysis only.

There were in fact some superficial trends in terms of clustering for the two categories, particularly this was described in the results section for Affiliation. In a binary problem with two categories (Affiliation and Recognition), there would logically be some discrimination between the two categories to allow visually evident clustering in the Affiliation

category. For Affiliation there was further clustering within this category, with two sets of pairs of questions within this category often grouped together. The Affiliation questions that were grouped together in pairs were "To socialize with other participants" and "To meet people" as one pair of questions. The second pair were the questions "To visit with friends" and "To share a group identity with other people". This result would imply two separate subsets of Social Motives questions. In the case of the questions "To socialize with other participants" and "To meet people", this would seem logical in terms of interpretation of the language used, as both these questions contain very similar activities.

CONCLUSION

It was demonstrated that t-SNE could be utilised to produce two-dimensional graphs to visualize the relationship between the different Social Motives category psychological motivation questions comprising the MOMS tool. Such two-dimensional scatter plots produced using t-SNE may assist in creating hypotheses about the relationships present between psychological constructs in such high-dimensional data. The general categorization of questions in the MOMS had commonalities with the groupings apparent in t-SNE graphs created across a range of learning rates. There were also some differences demonstrated in the t-SNE graphs.

There was more disparity in the data in terms of observed clustering between categories than expected, however visual inspection did confirm the presence of some limited patterns of clustering, which could be used to develop insights into clustering relationships within the data. One pattern was observed that was logical in terms of the underlying meaning in the language usage as well as the specific groupings of questions used by MOMS.

Acknowledgements

The authors appreciate the time taken by the 3,298 WMG masters athletes in completing the 56 questions in the MOMS survey tool. The authors also appreciate the assistance of Evan Wills in data collection using Limesurvey and the Sydney World Masters Games Organising Committee in approving the project.

Conflicts of interest: None

REFERENCES

1. Reaburn P, Dascombe B. Endurance performance in masters athletes. *European Review of Aging and Physical Activity*. 2008;5(1):31.
2. Beaumont A, Campbell A, Grace F, Sculthorpe N. Cardiac Response to Exercise in Normal Ageing: What Can We Learn from Masters Athletes? *Current Cardiology Review*. 2018;14(4):245-53.
3. Buyukyazi G. Differences in blood lipids and apolipoproteins between master athletes, recreational athletes and sedentary men. *The Journal of Sports Medicine and Physical Fitness*. 2005;45(1):112.
4. Climstein M, Walsh J, Debeliso M, Heazlewood T, Sevene T, Adams K. Cardiovascular risk profiles of world masters games participants. *The Journal of sports medicine and physical fitness*. 2018;58(4):489-96.
5. Fien S, Climstein M, Quilter C, Buckley G, Henwood T, Grigg J, et al. Anthropometric, physical function and general health markers of Masters athletes: a cross-sectional study. *PeerJ*. 2017;5:e3768.
6. Mckendry J, Breen L, Shad BJ, Greig CA. Muscle morphology and performance in master athletes: A systematic review and meta-analyses. *Ageing Research Reviews*. 2018;45:62-82.
7. Walsh J, Heazlewood, I.T., Climstein, M., DeBeliso, M., Adams, K.J., Burke, S., Kettunen, J., editor The Obesity Prevalence in Masters Athletes, a Comparison of all World Masters Games Sports. *Proc Pre-Olympic Conference: International Convention on Science, Education and Medicine in Sport*; 2012.
8. Walsh J, Heazlewood I, DeBeliso M, Climstein M. Psychological factors in competitive masters athletes in the context of body mass index. *Sport Science*. 2018;10(2):8.
9. Walsh J, Heazlewood IT, DeBeliso M, Climstein M. Assessment of motivations of masters athletes at the World Masters Games. *The Sport Journal*. 2018;20.
10. Walsh J, Climstein M, Heazlewood I, DeBeliso M, Kettunen J, Sevene T, et al. Reduced prevalence of smoking in masters football codes (rugby union, soccer and touch football). *Journal of Science and Medicine in Sport*. 2012;15:S134.
11. Heazlewood I, Walsh J, Climstein M, Adams K, Sevene T, DeBeliso M, editors. Differences in Participant Motivation Based on Category of Body Mass Index and Gender. *Singapore Conference of Applied Psychology*; 2016: Springer, Singapore.
12. Walsh J, Climstein M, Heazlewood IT, Burke S, Kettunen J, Adams K, et al. The loess regression relationship between age and BMI for both Sydney World Masters Games athletes and the Australian national population. *International Journal of Biological and Medical Sciences*. 2011;1(1):33.
13. Walsh J, Heazlewood I, Climstein M, Burke S, Adams K, DeBeliso M, et al. Body mass index for Australian athletes participating in rugby union, soccer and touch football at the World Masters Games. *Journal of the World Academy of Science, Engineering and Technology*. 2011;7(77):1119-23.
14. Walsh J, Climstein M, Burke S, Kettunen J, Heazlewood I, DeBeliso M, et al. Obesity prevalence for athletes participating in soccer at the World Masters Games. *International SportMed Journal*. 2012;13(2):76-84.
15. Walsh J, Heazlewood, I.T., Climstein, M., DeBeliso, M., Adams, K.J., Burke, S., Kettunen, J., editor The effect of body mass index on motivations for competition at the Sydney World Masters Games. *Proc Pre-Olympic Conference: International Convention on Science, Education and Medicine in Sport*; 2012.
16. Walsh J, Climstein M, Heazlewood IT, DeBeliso M, Adams K, Burke S, et al. Body mass index of masters basketball players. *Medicina Sportiva*. 2013;7:1700-5.
17. Walsh J, Climstein M, Heazlewood IT, Kettunen J, Burke S, Debeliso M, et al. Body mass index for athletes participating in swimming at the World Masters Games. *The Journal of Sports Medicine and Physical Fitness*. 2013;53(2):162-8.
18. Walsh J, Heazlewood IT, DeBeliso M, Climstein M. Comparison of obesity prevalence across 28 world masters games sports. *Sport Science*. 2018;11(1):30-6.
19. Climstein M, Walsh J, DeBeliso M, Heazlewood IT, Sevene T. CV Risk Profiles of WMG athletes PUBLISHED 2018. pdf. *The Journal of Sports Medicine and Physical Fitness*. 2018;58(4):489-96.
20. Heazlewood I, Walsh J, Climstein M, Adams K, Sevene T, DeBeliso M. Injury location, type and incidence of male and female athletes competing at the world masters games. *Journal of Science and Medicine in Sport*. 2017;20:e51-e2.
21. Walsh J, Climstein M, Heazlewood I, Adams K, DeBeliso M, Burke S, et al. Rugby union, soccer, touch football: Injury classification (masters athletes). *Journal of Science and Medicine in Sport*. 2011;14:e76.
22. Walsh J, Climstein M, Heazlewood I, Adams K, DeBeliso M, Burke S, et al. Masters athletes: Are they hurt more often?(rugby union, soccer and touch football). *Journal of Science and Medicine in Sport*. 2011;14:e76-e7.
23. Walsh J, Climstein M, Heazlewood IT, DeBeliso M, Kettunen J, Sevene TG, et al. Masters athletes: No evidence of increased incidence of injury in football code athletes. *Advances in Physical Education*. 2013;3(1):36-42.
24. Climstein M, Walsh J, Burke S, Adams K, DeBeliso M, Kuttunen J, et al. Physiological demographics of the Sydney World Masters Games competitors. *Journal of Science and Medicine in Sport*. 2011;14:e80.
25. Climstein M, Walsh J, Heazlewood I, DeBeliso M, Adams K, Sevene T, et al. Rowing, Soccer and Swimming: Injury differences (location, type, consequence) in preparation for the World Masters Games. *Journal of Science and Medicine in Sport*. 2012;15:S129.
26. Climstein M, Walsh J, DeBeliso M, Heazlewood T, Sevene T, Adams KJ. Cardiovascular risk profiles of world masters games participants. *The Journal of Sports Medicine and Physical Fitness*. 2018;58(4):489-96.
27. DeBeliso M, Climstein M, Adams K, Walsh J, Burke S, Heazlewood I, et al. North American medical and health history survey of 2009 Sydney World Masters Games participants. *Journal of Science and Medicine in Sport*. 2011;14:e79-e80.

28. DeBeliso M, Walsh J, Climstein M, Heazlewood I, Kettunen J, Sevene T, et al. World Masters Games: North American participant medical and health history survey. *The Sport Journal*. 2014;2014:1-17.
29. DeBeliso M, Walsh J, Heazlewood T, Sevene T, Adams KJ, Climstein M. Cardiovascular Risk Profiles Of World Masters Games Participants: 1023 Board# 202 May 31 200 PM-330 PM. *Medicine & Science in Sports & Exercise*. 2017;49(5S):277.
30. Climstein M, Burke S, Walsh J, Adams K, DeBeliso M, Heazlewood I, et al. Sydney 2009 World Masters Games: Participants medical and health history survey. *Journal of Science and Medicine in Sport*. 2010;13:e71.
31. Climstein M, Heazlewood IT, Walsh T, Walsh J, DeBeliso M, Sevene T, et al., editors. Non-optimal blood panels in World Masters Games participants. Proc ESSA conference 2012 (6th Exercise & Sports Science Australia Conference and Sports Dietitians Australia Update: Research to Practice); 2014.
32. Masters KS, Ogles BM, Jolton JA. The development of an instrument to measure motivation for marathon running: The Motivations of Marathoners Scales (MOMS). *Research Quarterly for Exercise and Sport*. 1993;64(2):134-43.
33. Marcus BH, Forsyth LH. *Motivating people to be physically active: Human Kinetics*; 2008.
34. Weinberg RS, Gould D. *Foundations of sport and exercise psychology*, 7E: Human Kinetics; 2018.
35. Heazlewood I, Walsh J, Climstein M, DeBeliso M, Adams K, Kettunen J, et al. The motivations of marathoners scales instrument for evaluating motivational factors in a variety of mainstream sports. *Journal of Science and Medicine in Sport*. 2012;15:S137.
36. Heazlewood IT, Walsh J, Climstein M, DeBeliso M, Adams KJ, Burke S, et al., editors. The Athlete Motivations to Participate in the 2010 Pan Pacific Masters Games. Proc International Convention on Science, Education and Medicine in Sport; 2012: SECC.
37. Heazlewood I, Walsh J, Climstein M, editors. Can participant motivation predict training frequency and training type in Male Masters Rugby players competing at the 2010 World Golden Oldies Rugby Festival? Singapore Conference on Applied Psychology 2018; 2018.
38. Brown CS, Masters KS, Huebschmann AG. Identifying motives of midlife black triathlete women using survey transformation to guide qualitative inquiry. *Journal of Cross Cultural Gerontology*. 2018;33(1):1-20.
39. Heazlewood T, Walsh J, editors. Data Mining: Applications of Neural Network Analysis in Exercise and Sport Science. Proceedings 8th International Conference on Computer Science Education: Innovation and Technology (CSEIT 2017); 2017.
40. Heazlewood IT, Walsh J, Climstein M, Kettunen J, Adams K, DeBeliso M, editors. A comparison of classification accuracy for gender using neural networks multilayer perceptron (MLP), radial basis function (RBF) procedures compared to discriminant function analysis and logistic regression based on nine sports psychological constr. In Proceedings of the 10th International Symposium on Computer Science in Sports (ISCSS); 2016: Springer, Cham.
41. Walsh J, Heazlewood T, Climstein M. Application of gradient boosted trees to gender prediction based on motivations of master athletes. *Model Assisted Statistics and Applications*. 2018;13(3):235-52.
42. Walsh J, Heazlewood IT, Climstein M. Regularized linear and gradient boosted ensemble methods to predict athletes' gender based on a survey of masters athletes. *Model Assisted Statistics and Applications*. 2019;14(1):47-64.
43. Havenar J, Lochbaum M. Differences in participation motives of first-time marathon finishers and pre-race dropouts. *Journal of Sport Behavior*. 2007;30(3):270.
44. Buning RJ, Walker M. Differentiating mass participant sport event consumers: traditional versus non-traditional events. *Sport Marketing Quarterly*. 2016;25(1):47.
45. Ruiz F, Zarauz Sancho A. Validation of the Spanish version of the Motivations of Marathoners Scales (MOMS). *Revista Latinoamericana de Psicología*. 2011;43(1):139-56.
46. Ogles BM, Masters KS. A typology of marathon runners based on cluster analysis of motivations. *Journal of Sport Behaviour*. 2003;26(1):69-85.
47. Adams KJ, DeBeliso M, Walsh J, Burke S, Heazlewood IT, Kettunen J, et al. Motivations to Participate in Sport at the Sydney 2009 World Masters Games: 3236Board# 199 June 4 8: 00 AM-9: 30 AM. *Medicine & Science in Sports & Exercise*. 2011;43(5):940.
48. Adams K, DeBeliso M, Walsh J, Burke S, Heazlewood I, Kettunen J, et al. Why do people participate in the World Masters Games? *Journal of Science and Medicine in Sport*. 2011;14:e82.
49. Adams K, Walsh J, Burke S, Heazlewood I, T., Kettunen J, DeBeliso M, Climstein M. Motivations to participate in sport at the 2010 Pan Pacific Games. *Official Journal of the American College of Sports Medicine*. 2012;44(5).
50. Heazlewood I, Walsh J, Climstein M, Burke S, Kettunen J, Adams K, et al. Sport psychological constructs related to participation in the 2009 World Masters Games. *World Academy of Science, Engineering and Technology*. 2011;7:2027-32.
51. Heazlewood I, Walsh J, Climstein M, Burke S, Kettunen J, Adams K, et al., editors. Athlete motivations to participate in the 2010 Pan Pacific Masters Games. ICSEMIS 2012 Pre-Olympic Conference: Sport Inspiring a Learning Legacy; 2012.
52. Heazlewood I, Walsh J, Climstein M, Adams K, Sevene T, DeBeliso M, et al. Gender differences in participant motivation in masters football at the 2010 Pan Pacific Masters Games. *International Research in Science and Soccer II: Routledge*; 2015. p. 236.
53. Heazlewood I, Walsh J, Climstein M, Adams K, Sevene T, DeBeliso M, editors. Participant Motivation Predicting Training Sessions and Training Type in Male and Female Athletes Competing at 2010 Pan Pacific Masters Games. Singapore Conference of Applied Psychology; 2016: Springer, Singapore.
54. Heazlewood IT, Walsh J, Climstein M, Adams K, Sevene T, DeBeliso M, et al., editors. Participant motivation: A comparison of male and female athletes competing at the 2009 World Masters Games. *Applied Psychology: Proceedings of the 2015 Asian Congress of Applied Psychology (ACAP 2015)*; 2016.
55. Sevene T, Adams K, Climstein M, Walsh J, Heazlewood I, DeBeliso M, et al. Are masters athletes primarily motivated by intrinsic or extrinsic factors? *Journal of Science and Medicine in Sport*. 2012;15:S357.
56. Heazlewood I, Walsh J, Climstein M. Re-evaluation of the factor structure of motivations of marathoners scales (MOMS). *Applied Psychology Readings: Springer*, Singapore; 2018. p. 57-71.
57. Maaten Lvd, Hinton G. Visualizing data using t-SNE. *Journal of Machine Learning Research*. 2008;9(Nov):2579-605.
58. Hinton GE, Roweis ST, editors. Stochastic neighbor embedding. *Advances in Neural Information Processing Systems*; 2003.
59. Kullback S, Leibler RA. On information and sufficiency. *The Annals of Mathematical Statistics*. 1951;22(1):79-86.
60. De Oliveira MF, Levkowitz H. From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics*. 2003;9(3):378-94.
61. Walsh J, Heazlewood IT, DeBeliso M, Climstein M. A profile of Sydney World Masters Games athletes: Health, injury and psychological indices. *Central European Journal of Sport Sciences and Medicine*. 2018;23(3):37-52.
62. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*. 2011;12(Oct):2825-30.
63. Scikitlearn. sklearn.manifold.TSNE 2018 [Available from: <http://scikitlearn.org/stable/modules/generated/sklearn.manifold.TSNE.html>].
64. Altman NS. An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician*. 1992;46(3):175-85.
65. Heazlewood IT, Walsh J, Climstein M, Burke S, Kettunen J, Adams KJ, et al., editors. Sport psychological constructs related to participation in the 2009 World Masters Games. Proc Seventh International Conference on Sport Medicine and Sport Science; 2011.

Creative Commons (CC) License-

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. (<http://creativecommons.org/licenses/by/4.0/>).