Number of trials necessary to achieve reliable change-of-direction measurement in amateur basketball players

Wing-Kai Lam1,2*, Yi Qu2, Fan Yang2, Rui-Ya Ma2, Zuo-liang Liu2 and Ben-Gang Yu2

Abstract: This study examined the inter-day reliability and repeatability of change-of-direction (COD) performance to determine the minimum number of measurement trials required to represent a stable athlete’s performance. Twenty-three university basketball players performed seven COD trials. Identical tests were performed two weeks later to evaluate inter-day reliability. Intra-class correlation coefficients were utilised to examine inter-day reliability for all data input methods. Separated ANOVA with repeated measures was performed to examine the trial effect. The ICC analysis indicated that data from the first trial exhibited the lowest reliability and that the fastest trial generally had higher reliability than the mean performance trial. The trial effect results revealed slower COD performance time in the first three trials than in later trials \( (p < 0.05) \), with no significant differences amongst later trials. These results suggested that at least four change-of-direction trials are required with respect to improving measurement reliability.

Subjects: Sport and Exercise Science; Sports Coaching; Physical Education; Research Methods

Keywords: athletic performance; agility; inter-trial reliability

ABOUT THE AUTHORS

Wing-Kai Lam is the head and director for LN Sports Science Research Center and also a visiting professor at the Department of Kinesiology, Shenyang Sport University, Shenyang, China; Li Ning Sports Science Research Center, Beijing, China. He completed his MPhil at Orthopaedics Surgery Dept and PhD at Institute of Human Performance, University of Hong Kong. He was an active basketball coach for athletes at various competition levels. He leads the footwear innovation and research activities in various court sports for his company. His research areas are ranging from biomechanics, footwear perception, sport injury to motor learning and performance.

The team actively researchers within the following areas: sport biomechanics, sport medicine, motor learning, sport training and computer stimulation. The information will be used to optimise sports products and to design various testing protocols which assess the performance across conditions.

PUBLIC INTEREST STATEMENT

Agility or change-of-direction performance assessment is one of the routine tests in sport and physical education research. It indicates an athlete’s baseline capability or measures relative changes between training sessions. However, using too few trials cannot represent an individual’s performance, while collecting too many trials may induce fatigue. This study determined the minimum required agility measurement trials to attain performance stability and identify the best data selection method for accurately representing agility performance in basketball. The present results indicated the best performance trials appear to be more reproducible across different test days than the mean performance trials. Additionally, the first trial demonstrated poor inter-day reliability. To establish change-of-direction protocols, at least four change-of-direction trials were required in order to make use of the average or best trial for reliable interpretation. Such information may assist researchers and coaches in the design of reliable agility protocols that minimise participants’ testing load and reduce unnecessary data processing time.
1. Introduction

Change-of-direction (COD) ability is a key performance measurement in various sports, such as basketball, badminton and football (Gamble, 2012) and is recognised as the ability to accelerate, decelerate and change direction rapidly (Gamble, 2012; Sheppard & Young, 2006). It has been commonly measured to indicate an athlete’s baseline capability or to measure relative changes between interventions when quantifying pre-/post-treatment training or intervention outcomes (Gamble, 2012; Hong, Lam, Wang, & Cheung, 2016; Liu, Wu, & Lam, 2017; Sheppard & Young, 2006). However, without an understanding of the reliability of COD manoeuvres, the accuracy of the athletic performance assessment is compromised. Moreover, measures which do not produce consistent results cannot be regarded as valid (Hopkins, 2000). The reliability is defined as the number of trials necessary for estimating a population parameter value (Hopkins, 2000) and thus the number of trials required for reliable task representation is an important methodological consideration in studies and assessments involving COD manoeuvres.

Previous studies determining minimum number of trials required to reliably represent athlete performance have primarily focused on running, walking, jumping and landing tasks (Bates, Osternig, Sawhill & James, 1983; DeVita & Bates, 1998; Hamill & McNiven, 1990; James, Herman, Dufek, & Bates, 2007; Oriwol, Milani, & Maiwald, 2012), but little attention was given to basketball-specific movements (e.g. COD tasks). In basketball step-off landing, James et al. (2007) reported at least four measurements which were required to obtain stable and reliable ground reaction force results; In basketball lay-up, Chua, Quek, and Kong (2016) found that a minimum of eight measurements were required to obtain stable and reliable foot plantar pressure variables. These findings suggest that the minimum number of trials would be specific to the movement tasks and thereby not applicable to other basketball COD tasks. Given that the prevalent use of COD tasks in the literature involves more complex movement coordination patterns and multiple repeated movements, an investigation of the minimum number of trials required to reliably represent this manoeuvre is warranted.

Performance studies on COD movements employing two or three trials found that inferior performances (i.e. lower speed or increased completion time) were observed in the first trial than in the second or third trial (Hexagon agility test, Beekhuizen, Davis, Kolber, & Cheng, 2009; T-agility test, Chaouachi et al., 2009; Illinois agility test, Hachana et al., 2013). These findings imply that interpretations based on an average of three trials may not represent an athlete’s performance particularly well. Using too few trials cannot accurately represent an individual’s performance (Devita & Bates, 1998; Oriwol et al., 2012); however, conducting too many trials with maximum effort may also induce fatigue (Amiri-Khorasani, Osman, & Yusof, 2010). Furthermore, even though other studies use the best performance trials as measurement outcomes, it is questionable that the best performance trials were obtained from different number of trials. Such information may assist researchers and coaches in the design of reliable protocols that minimise participants’ testing load and reduce unnecessary data processing time (Bates et al., 1983; Chua et al., 2016).

The present study sought to determine the minimum number of COD trials required to reflect performance stability by analysing trends over seven trials and inter-day reliabilities for different data input methods. Based on the previous findings, it is hypothesised that increasing number of trials would result in better stability of the COD measurements and thus reliable mean and/or best trials from different data input methods.

2. Methods

2.1. Participants

Twenty-three male university basketball players (age = 23.5 ± 1.7 years, height = 177.6 ± 4.5 cm, mass = 75.1 ± 6.7 kg) participated in this study. These players were all active participants in university-level competition and had an average playing experience of 7.2 ± 2.5 years. These athletes, instead of professional players, represented a wider population of amateur basketball players. All participants were right-leg dominant and had no lower extremity pain or injury in the past six
months prior to the start of the study and were required to refrain from any sporting activities in the past 24 h prior to the test. All participants underwent the second testing session at the same time of the day as the first testing session. Written consent was obtained from each participant and the test procedure was approved by the institutional review board.

2.2. COD task
The COD task included a sequence of backward shuffling, acceleration, deceleration, lateral shuffling and countermovement vertical jumps. The detailed procedure is described in Figure 1. All participants were instructed to begin with their back facing the court underneath the basket and to end with a maximum jump upon their return to the starting position. The elapsed time taken to complete the course was measured from the instance of movement initiation (movement 1) to the instance of take-off for the jump (movement 7), in accordance with the COD protocol of Liu et al. (2017).

2.3. Procedure
Prior to data collection, participants performed 10 minutes of individual warm-up and stretching. The participants were then asked to tighten the laces of their shoes in accordance with their personal habits for a basketball game. Three to five mandatory practice trials performed at a submaximal effort were used to familiarise participants with the COD protocol prior to data acquisition. Participants were instructed to complete the protocol as quickly as possible. The recording time, nearest to 0.01 s, was measured to indicate COD performance, as described by previous studies (Liu et al., 2017; Mclean, Lipfert, & van den Bogert, 2004). Additionally, the elapsed time was validated with video analysis at a capturing frequency of 300 Hz (Casio High Speed Exilim Ex-zr300, Casio, Tokyo) and had a high reliability between measurement tools (ICC = 0.96). A successful trial was defined as any trial without obvious slippage, discontinuity of movement or crossing of the legs during shuffling. Seven successful trials were collected based on the results of pilot testing. To minimise potential fatigue, two-minute rest periods between trials were required (Liu et al., 2017). To evaluate inter-day reliability, the identical test protocol was repeated two weeks later, which had a similar separate period to the previous studies (Lam, Sterzing, & Cheung, 2011, 2013).

2.4. Data analysis
Intra-class correlation coefficients [ICC(2,1)] for all data input methods were utilised to examine inter-day reliability. The reliability was classified as slight (>0.0 to ≤0.2), fair (>0.2 to ≤0.4), moderate (>0.4 to ≤0.6), substantial (>0.6 to ≤0.8) and almost perfect (>0.8 to ≤1.0), according to definitions by Portney and Watkins (2009). The data input selections included the mean and fastest trials obtained.
from the first two trials up to all seven trials; these selections have commonly been used to represent actual performance in many studies of COD and speed performance (Gamble, 2012; Moir, Button, Glaister, & Stone, 2004). Furthermore, separated ANOVA with repeated measures was performed to examine trial effect on COD performance time. Bonferroni-adjusted pairwise comparisons were performed post hoc where appropriate. The level of significance was set at 0.05.

3. Results

Observations of inter-day ICCs for data input methods revealed that the data from the first trial were the least reliable (Fair, Table 1). For the mean performance trials, the mean of the first two trials was moderate reliability while the mean values from all other data input selections were substantial reliability (Table 1). For the fastest performance trials, all data input selections were substantial reliability (Table 1). Furthermore, the inter-ICC was higher for the “fastest trial” data compared with the “mean trial” data.

On Day 1, the ANOVA of COD performance times (Figure 2) revealed a significant trial effect $F(6, 132) = 6.94, p < 0.001, \eta^2 = 0.24, \beta = 1.00)$. The post hoc analysis indicated a significantly slower performance time for the first and second trials compared with later trials (i.e. the third to seventh trials) ($p < 0.05$ for all comparisons). On Day 2, the ANOVA indicated a significant trial effect $F(6, 132) = 2.51, p < 0.05, \eta^2 = 0.10, \beta = 0.82)$. The post hoc analysis indicated a significantly slower performance time for the first three trials compared with the seventh trial ($p < 0.05$ for all comparisons).

| Table 1. Inter-day intra-class correlation coefficients (ICCs) and reliability level by various data input selections |
|---------------------------------------------------------------------------------------------------------------|
| **ICC** | **Reliability level** |
| The first trial | 0.312 | Fair |
| Mean of 1–2 trials | 0.548 | Moderate |
| Mean of 1–3 trials | 0.605 | Substantial |
| Mean of 1–4 trials | 0.616 | Substantial |
| Mean of 1–5 trials | 0.684 | Substantial |
| Mean of 1–6 trials | 0.685 | Substantial |
| Mean of all 7 trials | 0.715 | Substantial |
| Fastest trial of 1–2 trials | 0.648 | Substantial |
| Fastest trial of 1–3 trials | 0.678 | Substantial |
| Fastest trial of 1–4 trials | 0.692 | Substantial |
| Fastest trial of 1–5 trials | 0.730 | Substantial |
| Fastest trial of 1–6 trials | 0.708 | Substantial |
| Fastest trial of all 7 trials | 0.750 | Substantial |

Figure 2. Mean change-of-direction performance times across trials on days 1 and 2.

Notes: *, # and ^ denote significant differences ($p < 0.05$) compared with the first and second trials, respectively.
4. Discussion

COD performance assessment is an essential functional assessment in exercise science and sport conditioning research. Using too few trials cannot represent an individual’s performance (Devita & Bates, 1998; Oriwol et al., 2012), while collecting too many trials may induce fatigue (Amiri-Khorasani et al., 2010). The present study sought to determine the minimum number of trials required to accurately assess COD performance, thereby providing scientific guidelines for evaluating COD performance in basketball. This information may assist researchers and coaches in the design of reliable protocols that minimise participants’ testing load and reduce unnecessary data processing time (Bates et al., 1983; Chua et al., 2016). The ICC findings suggest that the first trial data were the least reliable (i.e. fair level of reliability) and that the average of the first two trials exhibited moderate inter-day reliability. These can provide insights to extract the necessary information/trials (i.e. cautious with the first two trials seems necessary) from the data-set for improving the reliability. Furthermore, the inter-day reliability was systematically better for the “fastest trial” data compared with the “mean trial” data, suggesting that selecting fastest/best trials would be recommended for COD measurements between sessions. Similar phenomenon would be found in the previous studies on other basketball assessments (Lam et al., 2013). From the physical measurement perspective, it is important to assure that COD abilities are measured using a reliable test protocol with data mining procedures to better understand the efficiency of training regimes.

The ANOVA findings suggest that the performance timing of the first and second trials was slower compared with later trials (i.e. the third to seventh trials) on Day 1, while performance timing of the first three trials was slower compared with the seventh trial on Day 2. The present findings are in line with the results from the studies on various COD movements (Hexagon agility test, Beekhuizen et al., 2009; T-agility test, Chaouachi et al., 2009; Illinois agility test, Hachana et al., 2013), which reported lower speed (or longer completion time) for the first trial compared with the second or third trial. The plausible explanation of the improving results (i.e. the first few trials) is that learning or adaptation processes could potentially have been involved across trials and may have varied across participants (Lam et al., 2013). Collectively, both inter-day reliability and ANOVA results suggest that at least four COD trials are required for taking the average or best trial for reliable interpretation. Such information can benefit researchers and coaches to develop the test protocol which minimises the load imposed on participants and reduce unnecessary data collection and processing time (Amiri-Khorasani et al., 2010). Future studies can further explore the performance stability in other COD movements.

When interpreting the results, it is important to consider several limitations in this study. Firstly, only male amateur players were recruited and hence our findings may not be generalised to female and elite players. It is expected that reliable data may be acquired in fewer trials with high-level athletes compared with low-level athletes because of their consistent movement control/performance. Secondly, left and right directional change performances were not addressed in this study, which may have affected the stability of the COD performance. However, this was assumed not to affect the between trials reliability evaluation investigated in the present study. Notably, other factors, including fatigue, practice, participant variability, playing position, time between testing sessions, environmental conditions and the appropriateness of difficulty levels for participants, could affect a test’s reliability when evaluating differences between interventions. Studying these confounding factors would improve the reliability of COD performance for a general population of basketball players.

To conclude, the best performance trials appear to be more reproducible across different test days than the mean performance trials. Additionally, the first trial demonstrated poor inter-day reliability. To establish COD protocols, at least four COD trials are required in order to make use of the average or best trial for reliable interpretation.
Funding
The authors received no direct funding for this research.

Author details
Wing-Kai Lam1,2 E-mail: gilbertlam@li-ning.com.cn
Yi Qu2 E-mail: quy@li-ning.com.cn
Fan Yang2 E-mail: yangfan6@li-ning.com.cn
Rui-Ya Ma2 E-mail: maruiya@li-ning.com.cn
Zuo-liang Liu2 E-mail: liuzhuliang@li-ning.com.cn
Ben-Gang Yu2 E-mail: yubengang@li-ning.com.cn

1 Department of Kinesiology, Shenyang Sport University, Shenyang, China.
2 Li Ning Sports Science Research Center, Beijing, China.

Citation information
Cite this article as: Number of trials necessary to achieve reliable change-of-direction measurement in amateur basketball players, Wing-Kai Lam, Yi Qu, Fan Yang, Rui-Ya Ma, Zuo-liang Liu & Ben-Gang Yu, Cogent Education (2018), 5: 1428386.

References
Amiri-Khorasani, M., Osman, N. A. A., & Yusof, A. (2010). Kinematics analysis: Number of trials necessary to achieve performance stability during soccer instep kicking. Journal of Human Kinetics, 23, 15–20.
Bates, B. T., Ostering, L. R., Sawhill, J. A., & James, S. L. (1983). An assessment of subject variability, subject-shoe interaction, and the evaluation of running shoes using ground reaction force data. Journal of Biomechanics, 16(3), 181–191. https://doi.org/10.1016/0021-9290(83)90125-2
Beekhuizen, K. S., Davis, M. D., Kolber, M. J., & Cheng, M. S. (2009). Test-retest reliability and minimal detectable change of the hexagon agility test. Journal of Strength and Conditioning Research, 23(7), 2167–2171. https://doi.org/10.1519/JSC.0b013e3181b43f0
Chaussac, A., Brughelli, M., Chamari, K., Levin, G. T., Abdelkrim, N., Laurencelle, L., & Costagna, C. (2009). Lower limb maximal dynamic strength and agility determinants in elite basketball players. Journal of Strength and Conditioning Research, 23(5), 1570–1577. https://doi.org/10.1519/JSC.0b013e3181e37f0
Chua, K., Quek, R. K. K., & Kong, P. W. (2016). Basketball lay-up – foot loading characteristics and the number of trials necessary to obtain stable plantar pressure variables. Sport Biomechanics, 31, 1–10.
Devita, P., & Bates, B. T. (1998). Intraday reliability of ground reaction force data. Human Movement Science, 7, 73–85.
Gamble, P. (2012). Training for sports speed and agility (pp. 1–6). London: Routledge.

Hachana, Y., Chao, L., Nabi, M. A., Attia, A., Moualhi, J., Farhat, N., & Elloumi, M. (2013). Test-retest reliability, criterion related validity and minimal detectable change of the Illinois agility test in male team sport athletes. Journal of Strength and Conditioning Research, 27(10), 1752–1759.
Hamill, J., & McNiven, S. L. (1990). Reliability of selected ground reaction force parameters during walking. Human Movement Science, 9, 117–131. https://doi.org/10.1016/0167-9457(90)90023-7
Hong, Y., Lam, W. K., Wang, S., & Cheung, J. T. (2016). Changes in comfort perception and direction change performance of badminton shoes with extensive usage time. Footwear Science, 8(1), 13–17. https://doi.org/10.1080/19424280.2015.1130753
Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. Sports Medicine, 30(1), 1–15. https://doi.org/10.2165/00007256-200030010-00001
James, C. R., Herman, J. A., Dufek, J. S., & Bates, B. T. (2007). Number of trials necessary to achieve performance stability of selected ground reaction force variables during landing. Journal of Sports & Science and Medicine, 6, 126–134.
Lam, W. K., Sterzing, T., & Cheung, J. T. (2011). Reliability of a basketball specific testing protocol for footwear fit and comfort perception. Footwear Science, 3(3), 151–158. https://doi.org/10.1080/19424280.2011.630680
Lam, W. K., Sterzing, T., & Cheung, J. T. (2013). Influence of protocol complexity on fit perception of basketball footwear. Footwear Science, 5, 155–163. https://doi.org/10.1080/19424280.2013.789933
Liu, H., Wu, Z., & Lam, W. K. (2017). Collar height and heel counter stiffness for ankle stability and athletic performance in basketball. Research in Sports Medicine, 25, 209–218. https://doi.org/10.1080/15438627.2017.123552
Mclean, S. G., Lipfert, S. W., & van den Bogert, A. J. (2004). Effect of gender and defensive opponent on the biomechanics of sidestep cutting. Journal of Biomechanics, 37(6), 73–85. https://doi.org/10.1016/j.jbiomech.2003.08.004
Moir, G., Button, C., Gloister, M., & Stone, M. H. (2004). Influence of familiarization on the reliability of vertical jump and acceleration sprinting performance in physically active men. Journal of Strength and Conditioning Research, 18(2), 276–280.
Oriwol, D., Milani, T. L., & Maiwald, C. (2012). Methodological issues associated with the mean value of repeated laboratory running measurements. Footwear Science, 4(3), 183–190. https://doi.org/10.1080/19424280.2012.676676
Portney, L. G., & Watkins, M. P. (2009). Foundations of clinical research: Applications to practice (pp. 585–618). Upper Saddle River, NJ: Pearson Prentice hall.
Sheppard, J. M., & Young, W. B. (2006). Agility literature review: Classification, training and testing. Journal of Sports Sciences, 24(9), 919–932. https://doi.org/10.1080/026041104000457109
