Two Sides of the Same Coin:
Do the Dual 'Types' of Passion Describe
Distinct Subgroups of Individuals?

Julia Moeller[a], Melanie Keiner[b], & Robert Grassinger[c]

[a] Yale Center for Emotional Intelligence, 340 Edwards Street, P.O. Box 208376, New Haven, CT 06520-8376, U.S.A.
[b] University of Erfurt, Germany
[c] University of Augsburg, Germany

Email addresses:

julia.moeller@yale.edu; melanie.keiner@uni-erfurt.de; robert.grassinger@phil.uni-augsburg.de

Abstract: The fact that harmonious and obsessive passion are called 'types of passion' (Vallerand, 2012; Vallerand et al., 2003) suggests that they describe subgroups of individuals. This study takes a person-oriented approach to examine groups of individuals with distinct profiles of harmonious passion (HP) and obsessive passion (OP). The aim of the study was to investigate whether HP and OP describe types in the sense of distinct groups of individuals with different passion profiles (intra-individual constellations), or if HP and OP are instead two constructs with distinct outcomes but possible alignment within individuals. Toward this aim, we analyzed four different samples ($N_{total} = 1149$) of German and Brazilian adolescents and young adults. Using HP and OP as classification indicators, we compared three different classification procedures in terms of the resulting groups. We found that the previously applied classification based on *z*-scores of HP and OP was misleading, because it classified individuals as 'obsessive' who had reported higher HP than OP when responding to the original response scale. Alternative classification based on raw scores and cluster analyses revealed that most individuals experienced either co-occurring high HP and OP, or co-occurring low HP and OP, whereas the assumed mainly obsessive group was not found in any of our four samples. The general degree of passion rather than the distinction between harmonious and obsessive individuals accounted for inter-individual differences in passion, and related constructs such as dispositional approach motivation, and affect. The findings were replicated in all four samples. This person-oriented approach gives novel and unique insights about inter-individual differences and intra-individual variation that can be easily overlooked or misunderstood in merely variable-oriented approaches, which prevail in prior research. Implications for the research on passion and pitfalls of *z*-scores in profile analyses are discussed.

Keywords: passion, person-oriented approach, profiles, standardization, motivation

Most passionate individuals strongly endorse the passionate items from both subscales, making it impossible to distinguish "pure" harmoniously passionate people (those who did not also endorse items from the obsessive subscale) from "pure" obsessively passion people without losing the majority of the sample to a mixed group. (Mageau et al., 2009, p. 613)

Introduction

Passion is a motivational construct that has been studied intensely during the last decade (for a review and for discussion of related constructs, see Moeller & Grassinger, 2014a). The prior research is dominated by studies about differences between the adaptive 'harmonious' passion (HP) and the maladaptive 'obsessive' passion (OP; Vallerand et

al., 2003). HP and OP are often referred to as 'types' of passion. Previous studies have interpreted the differences between both types as differences between harmonious versus obsessive individuals, stating that harmoniously passionate individuals make adaptive experiences of their passionate activity whereas obsessively passionate individuals make maladaptive experiences of their passionate activity (Vallerand et al., 2003; Vallerand, 2012).

Since these assumptions relied mostly on variable-oriented analyses, this study adopted a person-oriented approach to examine subgroups of individuals with distinct profiles of HP and OP. We apply different classification methods and discuss their advantages and disadvantages. Analyses of subgroups and intra-individual profiles lead to different conclusions than the variable-oriented approaches. Revising the classification methods applied in previous studies, we point out that group differences in regard to z-scores can be misleadingly interpreted, which has important implications not only for the research on passion, but for many other fields.

The dual model of passion

Passion describes the inclination of a person towards an activity that the person likes, finds important, identifies with and spends a lot of time and energy with (Vallerand et al., 2003). Passion helps to explain why individuals persist committed in effortful courses of action, such as intensive and persistent training (Vallerand et al., 2007). Although many different definitions of passion have been suggested in the psychological literature (Cardon, Wincent, Singh, & Drnovsek, 2009; Fredricks, Alfeld, & Eccles, 2010; Frijda, 2007; Moeller & Grassinger, 2014a; Renzulli, Köhler, & Fogarty, 2006; Solomon, 1993; Sternberg, 1997), the dual model of passion from Vallerand et al. (2003) is the most pervasive in the field.

This model states that passionate individuals identify with their passionate activity, and that a voluntary, autonomous internalization of the activity leads to harmonious passion, whereas the internalization due to internal or external constraints leads to obsessive passion (Mageau et al., 2009). The previous research has mostly focused on the distinction between HP and OP and their differences regarding predictors, correlates, and outcomes. The components of the dual model of passion are measured with the passion scale (Vallerand et al., 2003), which comprises one subscale tapping general criteria of passion to distinguish between non-passionate and passionate individuals (example item: "This activity is important for me"), a second subscale assessing HP (example item: "This activity is in harmony with the other activities in my life.") and a third subscale assessing OP (example item: "The urge is so strong. I can't help myself from doing this activity."). HP typically predicts and correlates with positive experiences such as psychological and physical well-being (e.g, Vallerand 2012), positive affect during and after activity engagement (Philippe, Vallerand, Houlfort, Lavigne, & Donahue, 2010), and maintained control over the activity (Vallerand, et al., 2003; Mageau, Vallerand, Rousseau, Ratelle, & Provencher, 2005). In contrast, OP often correlates with maladaptive experiences such as loss of control, symptoms of dependency (Philippe & Vallerand, 2007), negative affect during activity engagement and when the individual is prevented from doing the activity (Vallerand et al., 2003), and ill-being (Rip, Fortin, & Vallerand, 2006; Vallerand et al., 2003).

Are 'types of passion' types in the sense of subgroups of individuals?

Several researchers interpret these findings as differences between distinct groups/profiles in HP and OP, assuming that individuals are either predominantly harmoniously or predominantly obsessively passionate, as seen in statements like the following: "People with an HP are able to decide to terminate the relationship with the activity if they decide it has become a negative factor in their life." (Vallerand et al., 2003, p. 758), and "people with an obsessive passion can thus find themselves in the position of experiencing an uncontrollable urge to partake in the activity they view as important and enjoyable." (Vallerand, 2012, p. 3). The statement that harmonious individuals were happier and better adapted (Vallerand et al., 2012) suggests that the findings regarding HP and OP not only describe co-variation among the studied population, but the experiences of particular individuals.

However, previous studies rarely examined differences between groups of harmonious versus obsessive individuals but rather analyzed correlations or regressions between HP, OP, and relevant predictors, correlates, and outcomes across individuals. While this helps to understand the overall relation between the measures of HP, OP, and their different correlates, it has been criticized that the analyses of co-variation on the level of a population often does not describe the experiences of real individuals (Bergman & Magnusson, 1997; Molenaar, 2004; Reizle, 2013). A population might consist of distinct subgroups of individuals, in each of which the relation between the studied variables might be different than the overall correlation suggests. For instance, the negative correlation between school engagement and school burnout hides a substantial subgroup of individuals who experience high levels of both engagement and burnout (Tuominen-Soini & Salmela-Aro, 2014; Salmela-Aro, Moeller, Schneider, Spicer & Lavonen, 2015). Moreover, paths between multiple variables that seem significant and meaningful in path analyses do not necessarily describe the individual experience, because the individuals that drive the correlation between variable A and B are not necessarily the same that drive the correlation between variable B and C (see Reizle, 2013).

The conditions under which general trends like correlations hold true on the level of individuals (and not only on

the level of a population) are strict and rarely met (Molenaar & Campbell, 2009). Consequently, the well-studied fact that HP and OP have different correlates, predictors and outcomes (despite their substantial correlation with each other) neither indicates that individuals experience either a predominantly harmonious or a predominantly obsessive passion, nor that experiencing OP is necessarily accompanied by low HP, low positive affect, low control over the activity, etc. Similarly, it might be that despite their different correlates, HP and OP are experienced *together*, rather than occurring in profiles of either predominating HP, or predominating OP. It might also be that the relation between the items of HP and OP differs across subgroups of individuals (e.g., hobby gardeners versus Olympic athletes in intense pre-competition training), across time within individuals, or across activities that differ in how demanding or competitive they are. For instance, an activity might become more difficult to harmonize with other activities when the passionate individual has young children to take care of, or when a grant proposal deadline is approaching. Factors, correlations between factors, and item-factor loadings, might change across time or across groups of individuals (for a detailed discussion, see the concept of dimensional identity; von Eye & Bergman, 2003).

The fact that HP and OP are called 'types of passion' (Vallerand, 2012; Vallerand et al., 2003) suggests that they describe subgroups of individuals. In the psychological literature, the term 'types' often describes homogeneous groups of individuals, which are characterized by a specific *profile* in the variables that are of interest for the current analysis. Individuals of the same type display similar profiles and differ substantially from profiles of individuals belonging to other types. Examples of psychological typologies are the classification of criminal offenders into differrent homogeneous groups with different recidivism risk and different needs for treatment (e.g. Simourd et al., 1994), and the classification of individuals with psychopathological symptoms according to the onset of the problems and the prognosis of their development (e.g. Millon et al., 1996). The classic methods to classify individuals with similar profiles into homogeneous groups and to identify the distinct groups within a sample are cluster analytical approaches.

Classification of passion types in the previous literature

In contrast to classic typologies, few studies that labeled HP and OP 'types' really examined groups of individuals. Exceptions are the studies from Mageau et al. (2009), Philippe, Vallerand, & Lavigne (2009), and Vallerand and Houlfort (2003), who classified the individuals in their samples into three different groups: (1) non-passionate, (2) mainly harmonious and (3) mainly obsessive individuals. To distinguish passionate from non-passionate individuals, the authors used between three and four of the criteria for general passion (liking, finding important, investing resources and labelling the activity as a passion). Individuals with a score above the scale midpoint were defined as passionate, and all others as non-passionate. Then the passionate individuals were subdivided into the harmonious and the obsessive group by the following procedure: The subscales for HP and OP were z-standardized across all individuals (including non-passionate), and passionate individuals with higher z-scores for HP than OP were classified as 'harmonious', while passionate individuals with higher z-scores in OP than HP were classified as 'obsessive'. We will call this *the z-score based grouping procedure* below.

These studies found the expected group differences. For instance, individuals who experienced stronger autonomy were more likely to be classified at later time points as 'harmonious' than as 'obsessive', while in contrast individuals who strongly valued activity specialization and strongly identified with the activity were more likely classified as 'obsessive' at later time points (Mageau et al., 2009). Also, the 'harmonious' group reported higher well-being than the 'obsessive' group (Philippe et al., 2009).

To our knowledge, there is only one study using cluster analysis on the dual model passion scale, conducted by Wang, Khoo, Liu, and Divaharan (2008). The authors used the scores of HP and OP as indicators in a hierarchical cluster analysis and found unexpectedly that harmonious and obsessive levels were aligned in the resulting clusters, meaning that participants in the sample scored either both high, both moderate, or both low. There was no mainly obsessive group. The generalizability of these findings was limited because the study was conducted in only one activity context (digital gaming). Therefore replications are required.

Some reflections on differences and possible misinterpretations of z-scores and raw-score classifications

The z-score-based classification procedure applied by Mageau et al. (2009), Philippe et al. (2009), and Vallerand and Houlfort (2003), differs from classic cluster analyses in that it forces individuals into theoretically assumed groups using theory-based cut-offs, whereas cluster analyses estimate the number and variable profiles of the groups inductively, based on, for example, log likelihood or Euclidean distances between individual observations, and groups of observations. Unlike cluster analyses, the z-score-based classification procedure does not take into account the size of the difference between an individual's HP and OP scores, meaning an individual with zHP=1.000 and zOP=0.999 would be classified as 'harmonious', even though the difference of 0.001 between HP and OP might be uninformative. Also, the classification into passionate and non-passionate individuals based on the scale midpoint of the passion criteria subscale is relatively arbitrary, given that this cut-off is not based on empirical group differences and

varied between previous studies (Mageau et al., 2009; Philippe et al., 2009).

Further, in contrast to cluster analyses, which may detect all relevant types in the data, the z-score based classification procedure risks overlooking groups of individuals with other than the expected profiles. For instances, there might be a group with high scores in both HP and OP (see e.g. Wang & Yang, 2007), which cannot be detected with the z-score based procedure unless an individual's HP and OP z-scores are exactly the same. Moreover, the interpretation of the groups created with the z-score-based procedure might be misleading, given that both the 'harmonious' and the 'obsessive' group displayed higher raw scores of HP than OP (Mageau et al. 2009; Wang & Yang, 2007). One reason for that is the fact that the raw mean score of HP is generally higher than the raw mean score of OP (Mageau et al., 2009). Therefore, a high sample rank in OP (high OP z-score) can be at the same or lower level on the original response scale than a moderate rank (z-score) in HP.

There are several potential problems with z-scores: First, raw scores reflect the original frame of reference and an individual's acceptance versus rejection of an item much better than z-scores do. Second, raw scores provide information about absolute distances between two measurements, whereas z-scores obfuscate this information. We explain these problems in the following:

Reference Frame. The typical instruction for the dual model passion scale is "While thinking of your favorite activity and using the scale below, please indicate your level of agreement with each item" (Vallerand et al., 2003[1]). The response scale for both HP and OP is: 1 = *not agree at all*, 2 = *very slightly agree*, 3 = *slightly agree*, 4 = *moderately agree*, 5 = *mostly agree*, 6 = *strongly agree*, 7 = *very strongly agree*. The original response scale is the reference frame that an individual uses to report how strongly he or she agrees to a given passion item. The order of items of both scales is mixed, which suggests that individuals really used the same reference frame when rating their agreement to these items. It is important to notice that any score below the scale midpoint (= 4) probably means that an individual disagrees more than agrees with the item statement. Z-standardized scores reflect a different reference frame: A z-score reflects the rank of an individual's answers in relation to the answers of all other individuals in the same sample. If a variable has a very low acceptance among the sample (= high item difficulty), then a z-score above the mean can still reflect that the person rather denied most OP items, just less strongly than others denied these items (see e.g. the OP mean score in our studies 1 and 4 below).

Raw scores reflect the degree to which an individual affirms or rejects item statements, whereas this information gets lost in z-scores. Therefore, it is of questionable validity

to interpret an OP score of 3.8 as "high OP", because in the original frame of reference this score reflected more rejection than acceptance of the OP items. However, because of the low mean score of OP (see Table 2), this low raw score would be interpreted as a comparably high z-score in many samples. In contrast, because of the rather high average of HP, an individual's score of 5.0 would be rated as general acceptance of HP items based on raw scores, but as a "rather low" score based on z-scores, because 5.0 is below the sample average (in all our four samples, see Table 2). Paradoxically, the sample-level mean difference between overall high HP and overall low OP has the consequence that an individual can rate most OP items with "slightly agree/ disagree", and most HP items with "mostly agree", and might still end up with a higher z-score for OP than for HP. Another person might rate all HP and OP items with "4" (meaning rather indecisive), and still get a higher z-score for OP than HP. This makes it impossible to interpret the psychological meaning of z-score based OP-HP-profiles.

Absolute distances. Other information lost in z-scores is the absolute distance between two measurements. Imagine you want to plot the mean-level difference between two groups (see Figure 1). Using raw scores shows the absolute distance between the groups if the complete response scale is shown (and a truncated Y-axis should be avoided because it is considered misleading). In contrast, if z-scores are displayed, then in most cases the resulting graph only shows the part of the range of possible answers where variance (= group differences) was found. That equals zooming into the picture, and makes eventually small differences look big. This is illustrated in Figure 1, which compares the raw and z-scores of the groups created with the z-score-based grouping procedure in sample 1.

In sum, very few studies on passion have tried to describe subgroups of individuals. Most studies have described only the co-variation of HP, OP, and their correlates, predictors and outcomes on the population level. The few attempts to describe individual profiles of HP and OP so far have not supported the assumed duality of mainly harmoniously passionate individuals versus mainly obsessively passionate individuals. However, this requires systematic replication.

The present research: Do the 'types of passion' describe distinct subgroups of individuals?

This present study investigated whether the well-established duality of HP versus OP describes distinct groups of mainly harmonious versus mainly obsessive individuals. For this purpose we first examined with our own data the implications and possible interpretations of the z-score based grouping procedure that previous studies applied to distinguish such groups. In a second step, across four samples we examined classifications based on raw scores and conducted cluster analyses to explore groups and their profiles of HP and OP.

[1] See also
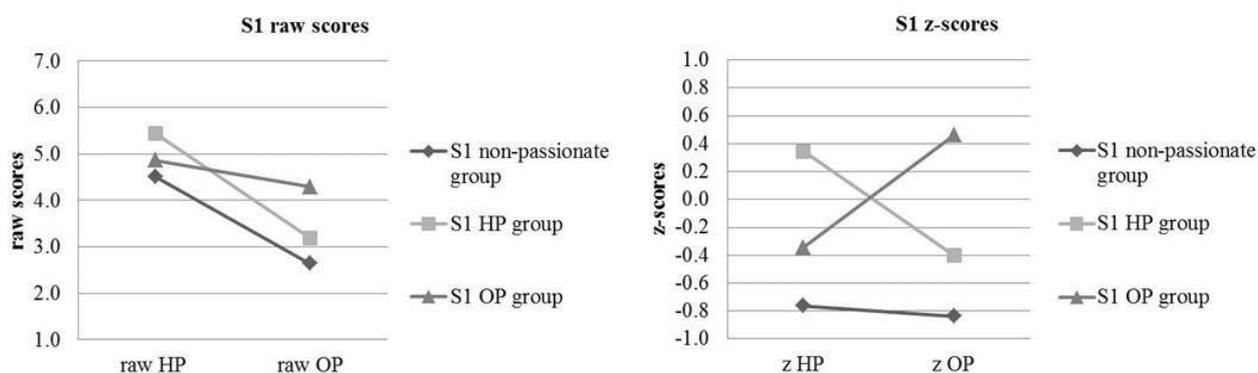http://www.er.uqam.ca/nobel/r26710/LRCS/echelles_en.htm

Figure 1: Illustrating the loss of information about absolute distances through z-standardization: Raw scores (left) and z-scores (right) of HP and OP in study 1.

Research questions

1) What characterizes the groups created with the z-score-based classification procedure?

   1a) Do individuals who are labeled "obsessive" according to the z-score-based procedure display higher raw-scores in OP than HP?

   1b) Do individuals who are labeled "harmonious" according to the z-score-based procedure display higher raw-scores in HP than OP?

We expected to find some misleadingly classified individuals (labeled 'obsessive' by the z-score-based procedure, but experiencing higher HP than OP according to their original answers in the questionnaire), because Mageau et al. (2009) and Wang and Yang (2007) had found the trend that individuals generally report higher HP than OP raw scores across all groups. Possible misinterpretations were examined in more detail.

2) Do alternative classification procedures such as grouping based on raw scores and cluster analysis also produce the expected groups of mainly harmonious and mainly obsessive individuals?

   2a) The raw-score based classification procedure forces individuals into either the mainly harmonious or the mainly obsessive group. Because of the generally higher levels of raw HP than OP, we expected that this procedure leads to a large group of 'mainly harmonious' and a rather small group of 'mainly obsessive' individuals.

   2b) With cluster analyses, we expected to find groups with generally high or low levels in both HP and OP, as found previously by Wang et al. (2008).

Methods

Samples

Passion and related constructs were assessed in four different samples, which were analyzed separately in order to test the replicability of the findings. Two samples of individuals responded to the passion items with reference to their most important favorite activity (sample one and two) and two samples responded with reference to particular leisure activities that they were practicing (sample three: leisure soccer; sample four: leisure dance and martial arts). These samples were selected because they resembled much the samples used in prior studies about the dual model of passion (in study 1 and 2, our assessment procedure was very similar to that of Vallerand et al., 2003; study 1, and in the studies 3 and 4, we used similar samples to those of Vallerand et al., 2003, study 2; Vallerand, Mageau, et al., 2008, study 1; and Mageau et al., 2009, study 2), so that our findings should be generalizable to what we would expect to find in the populations studied before by other researchers. Moreover, we selected these samples to check whether our findings would be replicated across different domains of leisure activities, different age groups (adolescents to mainly young adults), and different nationalities (Germany in the studies one, three and four; Brazil in study two).

Table 1 gives an overview about the samples and applied measures, which are described in detail below. Detailed information about the measures' reliabilities, validities, psychometric properties (item difficulty, skewness, curtosis, item-factor loadings, etc.), inter-correlations between subscales, and exact item wordings were documented by Moeller & Grassinger (2014b), and Moeller, Keiner, & Wächter (2014).

Table 1
Samples and Assessment

Sample	Assessment	N	Instruments
1	German undergraduate students (Universities of Erfurt and Augsburg),	471	Dual model passion scale [a], com.pass scale [b], Affect scale[c]
2	Brazilian adolescents in two private schools in Rio de Janeiro	274	Dual model passion scale [a], com.pass scale [b], sport commitment scale[d]
3	German adolescent soccer player, online survey and paper-&-pencil questionnaires	278	Dual model passion scale, com.pass scale [a], sport commitment scale[d], BAS[e], Affect scale: PANAS[f]
4	German leisure Dancers (149), martial arts practitioners (95)	244	Dual model passion scale[a], com.pass scale [b], BAS[e], Dependency measure EAI[g]

Note. [a] = Dual Model Passion Scale (Vallerand et al., 2003), [b] = com.pass scale (Moeller & Grassinger, 2014b); [c] = Affect Scale (Kessler & Staudinger, 2009); [d] = Sport Commitment Scale (Sousa et al., 2008); [e] = Behavioral Activation, subscale of the BIS/BAS scale (Carver & White, 1994; Strobel, Beauducel, Debener, & Brocke, 2001); [f] = Positive Negative Affect Schedule (Krohne, Egloff, Kohlmann, & Tausch, 1996); [g] = Exercise Addiction Inventory (Terry, Szabo, & Griffiths, 2004).

Participants of sample one

353 undergraduate students from the University of Erfurt (Germany) responded to questions in the lecture rooms between 2010 and 2013. Four participants were excluded from further analyses for not describing any passionate activity. The remaining 349 students were on average 23.63 years old (*SD* = 3.22), 67.0% were female. All participants studied psychology or a pedagogic major subject.

Participants of sample two

274 Brazilian private school students of two institutions were recruited in Rio de Janeiro in August 2012. The participants were on average 17.31 years old, 53.7% were female. The schools in which the data collection took place were two affiliated institutes of the private school "Intellectus" in the districts Méier and Catete in Rio de Janeiro. The participants attended the ninth grade (10.6%), tenth grade (26.3%), eleventh grade (25.9%), or curses in preparation for the university entrance exams (37.2%).

Participants of sample three

278 adolescent soccer players in the German federal state of Thuringia were surveyed in 2013. They were on average 14.87 years old. 99.1% were male. The sample was conducted in two steps. First, weblinks to the online survey were sent via mass e-mail to all soccer clubs in the German Federal state of Thuringia by the Thuringian Soccer Association (Thüringer Fußball- Verband, TFV). Together with that

e-mail, the clubs also received flyers and posters for the announcement of the study. All trainers were asked to announce the study with the help of these materials and to motivate all youth soccer players between 13 and 18 years to participate in the study. 48 individuals filled out the online questionnaires, two of them were excluded from further analyses because they were much older (> 27 years) than the target sample.

In order to achieve a larger sample size, we then contacted all Thuringian soccer clubs that trained competitive youth teams via telephone and e-mail for survey appointments. 17 clubs were then visited, and 234 participants filled out paper and pencil questionnaires during their soccer training lessons.

Participants of sample four

Sample four consisted of 244 individuals (128 leisure dancers and 95 martial art practitioners), who were surveyed with paper and pencil questionnaires in their dancing schools and training gyms. The two groups were combined to one sample in order to allow for analyses that require larger sample sizes.

Among the dancers, 56.6% were females, the mean age was 26.85 years (SD = 13.0). On average, the participants had practiced dance for 7.75 years. Among the martial arts practitioners, 36.6 % were females, the mean age was 33.67 years (SD = 14.36). On average, the participants had practiced martial arts for 9.33 years.

Table 2

Descriptive statistics and internal consistencies for the assessed scales

	Sample 1			Sample 2			Sample 3			Sample 4		
	M	*SD*	*α*	*M*	*SD*	*α*	*M*	*SD*	*α*	*M*	*SD*	*α*
Harmonious Passion [a]	5.15	.849	.73	5.29	.993	.69	5.48	.978	.89	5.35	.845	.73
Obsessive Passion [a]	3.70	1.266	.85	4.41	1.476	.87	4.34	1.540	.93	3.79	1.291	.85
General Passion criteria [a]	5.95	.722	.58	5.74	.905	.50	6.13	.876	.85	6.00	.991	.84
General Passion [b]	4.45	.734	.91	4.63	.745	.88	4.74	.715	.96	4.50	.662	.91
Commitment [c]				4.24	.595	.74	4.22	.699	.83			
Dispositional approach motivation [d]							3.30	.417	.86	3.10	.375	.76
Positive Affect [e]	4.00	.526	.75				4.16	.647	.91			
Negative Affect [e]	1.47	.423	.75				2.35	.997	.92			
Dependency [f]										3.14	.761	.73

Note. [a] = Dual Model Passion Scale (Vallerand et al., 2003), [b] = com.pass scale (Moeller & Grassinger, 2014b), [c] = Sport Commitment Scale (Sousa et al., 2008), [d] = BIS/BAS scale (Carver & White, 1994; Strobel, Beauducel, Debener, & Brocke, 2001), [e] = Positive Negative Affect Schedule (Krohne, Egloff, Kohlmann, & Tausch, 1996), [f] = Exercise Addiction Inventory (Terry, Szabo, & Griffiths, 2004).

Measures and measurement properties

In all four studies we used two different passion scales (Moeller & Grassinger, 2014b; Vallerand et al., 2003) in order to ensure the generalizability of the findings across specific measures. The reliabilities and validities of the passion scales were found to be satisfying in all four studies (for details, see Table 2 for Cronbach's α, and Moeller & Grassinger, 2014b, Moeller et al., 2014).

Measures of sample one

Passion was assessed with the Dual Model Passion Scale (Vallerand et al., 2003) and the com.pass scale (Moeller & Grassinger, 2014b). All constructs were measured with self-report questionnaires in German language. The students were instructed to name their three favorite activities, to rank these activities by importance, and to answer to all following questions in regard to the first ranked activity.

The Dual Model Passion Scale (Vallerand et al., 2003) consists of 18 items of which seven items are indicators for harmonious passion (α = .73), seven items are indicators of obsessive passion (α = .85), and four items are general passion criteria (α = .58). The response scale ranged from 1 = *do not agree at all* to 7 = *very strongly agree*. The German version of the scale was obtained via iterative forth- and backwards translation and subsequent discussion.

The concurrent validity of the German version of the Dual model passion scale was confirmed in this sample by the findings of significant moderate correlations between the dual model subscales and other measures of passion and commitment, the positive correlation of HP with high arousal positive affect, and the negative correlation of HP with low arousal negative affect (Moeller & Grassinger, 2014b; see also Moeller et al., 2014). These analyses also revealed that the EFA showed the expected factor structure, but CFA indicated below-optimal model fit due to residual

correlations, which is in line with previous findings (e.g. Marsh et al., 2013).

To validate the findings, all participants filled out the second passion measure, the com.pass scale (Moeller & Grassinger, 2014b), comprising 21 items with the four subscales Continuous Intent to Engage (six items), Identification (six items), Long-term Goals (five items), and Desire (four items). All items were answered on a 6-point scale with a range from 1 = *do not agree at all* to 6 = *very strongly agree*. The passion score is computed as the means score of the subscales Desire and Commitment, the latter consisting of the subscales Intent, Identification, and Goals.

A subsample of 206 students from the University of Erfurt also filled out an affect scale (Kessler & Staudinger, 2009), which is an adaptation of the Positive Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988). This affect scale assesses with four four-item subscales high arousal positive affect, low arousal positive affect, high arousal negative affect, and low arousal negative affect. To compare this measure with the PANAS affect measure applied in study three, we disregarded the affective intensity and summarized the items to one subscale of positive affect (α = .75) and one subscale of negative affect (α = .75).

Measures of sample two

The Brazilian students were surveyed in classes with the same procedure applied in study 1. In this sample, we used Portuguese versions of the above-described two passion scales. Both passion scales were independently translated into Portuguese by two Brazilian translators (one junior and one senior lawyer, both fluent in English with TOEFL scores > 100 and German skills above level A2 of the European reference frame for languages). The translators were presented with two equivalent versions of both passion scales (a German and an English version) and were asked to translate the scale into Portuguese. The first author then translated the scales back into German and discussed ambiguous transla-

tions with the translators in order to find the most equivalent Portuguese expression. The reliabilities of the dual model passion subscales were acceptable (see Table 2). Their validity was indicated by the expected correlations to other measures of passion and commitment (see Moeller & Grassinger, 2014b).

To validate the findings, commitment was measured with the commitment subscale of the Sport Commitment Questionnaire (SCQ; Scanlan et al., 2003), using the Portuguese version (Sousa et al., 2008), which according to our translators did not require language changes for being applied in Brazil. The SCQ consists of six subscales, one for the assessment of sport commitment as such and five subscales for the assessment of predictors of sport commitment. Since we were not interested in the predictors, we only assessed the sport commitment subscale. This scale consists of six items all of which tap the commitment towards the relevant activity, for example the Portuguese translation of the item "I am determined to play soccer in the next season". The response scale ranges from 1 = *not at all* to 5 = *very much*. In order to assess the commitment towards the individuals' favorite activities, the item wording was adapted to refer to activities in general (for Cronbach's Alpha see table 2).

Measures of sample three

As before, passion was measured with the Dual Model Passion Scale (Vallerand et al. (2003) and with the com.pass scale. In contrast to the samples one and two, ten new items were added to the com.pass scale (for details regarding the new items, see Moeller & Grassinger, 2014b). The reliabilities of the dual model passion subscales were good for HP and OP (see Table 2). The subscales' validity was indicated by the expected correlations between HP and OP to other measures of passion and commitment (see Moeller & Grassinger, 2014b).

As in sample two, sport commitment was assessed with the sport commitment scale. Since the latest publication of this scale was the Portuguese scale version from Sousa et al. (2008), we translated this version from Portuguese into German. The item wording was adapted to the activity 'soccer'.

Dispositional approach motivation was assessed with the BAS (behavioral activation system) subscale of the BIS/BAS scale (Carver and White, 1994), using the German translation from Strobel et al. (2001). The response scale ranges from 1 = *very true for me* to 4 = *very false for me,* and was reversed to match the polarity of the other scales, with low scores reflecting low degrees in dispositional approach motivation. While the original English version of the BIS/BAS distinguishes between three facets of BAS (reward, drive, and fun seeking), Strobel et al. (2001) found that a solution with a global BAS scale fitted their (German) data best. Therefore we report results for the global BAS score.

Dispositional affect in regard to soccer was assessed with the Positive Negative Affect Schedule (PANAS; German translation by Krohne et al., 1996), which taps positive

affect and negative affect each with ten items. The participants were asked how they had felt when playing soccer during the previous year. The response scale ranged from 1 = *not at all to* 5 = *very strongly*.

Measures of sample four

Again, passion was measured with the Dual Model Passion Scale and the com.pass scale (21 item version). These two scales again were moderately to highly inter-correlated, which was interpreted as indicator of concurrent validity. The reliabilities of the dual model passion subscales were satisfying for HP and OP (see Table 2). Their validity was indicated for instance by the expected positive correlation of OP with symptoms of dependency (see Moeller et al., 2014).

Behavioral approach motivation was assessed as in sample three with the BAS subscale of the BIS/BAS scale (Strobel et al., 2001).

Feeling dependent of dance and martial arts was assessed with the 6-item Exercise Addiction Inventory (EAI, Terry et al., 2004). The scale was translated in a forth-and-back translation process by students of translation (5th semester) and by native English speakers. The response scale ranges from 1 = *strongly disagree* to 5 = *strongly agree.* The internal consistency of this scale was acceptable (see Table 2).

Analyses

Research question 1: Do individuals who are labelled "obsessive" according to the z-score-based procedure display higher raw-scores in OP than HP, and vice versa?

First, we applied the same procedure as Mageau et al. (2009); Philippe et al. (2009) and Vallerand and Houlfort (2003) to categorize the individuals of each sample into a non-passionate group, a 'mainly harmonious' group, and a 'mainly obsessive' group. For this purpose we first standardized the variables HP and OP across all individuals within each sample. Then we used the scale midpoint of the general passion criteria scale to distinguish passionate individuals (mean score 4 or higher) from non-passionate individuals (mean score below 4). [2] Finally, we classified the passionate individuals into 'mainly harmonious' (zHP > zOP) and 'mainly obsessive' (zOP > zHP), and examined whether there were individuals with similar levels of (standardized) HP and OP.

In a second step, we analyzed and compared the z-scores and raw scores on HP and OP of the groups created with the z-score-based grouping procedure. Third, we analyzed how many individuals were classified as 'mainly obsessive'

---

[2] This cut-off was used by Mageau et al. (2009) and Vallerand & Houlfort (2003), whereas Philippe et al. (2009) set the cut-off at a mean score of 5 on the same scale.

with the *z*-score-based classification procedure, despite higher raw scores on HP than OP.

Research question 2: Do alternative classification procedures such as grouping based on raw scores and cluster analysis also produce the expected groups of mainly harmonious and mainly obsessive individuals?

Next, to examine how many individuals report higher HP than OP in their answers to the original response scale, we applied a similar classification procedure as described above, but used raw scores to classify individuals into a 'mainly harmonious' group (raw HP > raw OP), a 'mainly obsessive' group (raw OP > raw HP), and an equally harmonious and obsessive group (HP = OP). Then we compared how many individuals are classified as 'mainly harmonious' and 'mainly obsessive' with the z-score-based classification procedure versus the raw-score-based classification procedure.

Then, in order to visualize the multivariate distribution of scores, we examined the scatter plots for the participants' raw scores of HP and OP. Finally, to account for the size of distances between HP and OP, we examined subgroups in terms of each individual's profile of HP and OP by conducting cluster analyses. For this purpose we calculated in each of the four samples a two-step cluster analysis with SPSS (version 22). As indicators we used the scores of HP and OP[3]. This procedure suggests an optimal number of clusters based on Schwarz' Bayesian Criterion (BIC), the ratio of change in BIC from one model to another, and the ratio of change in distance measures. The quality of this final solution is evaluated based on the silhouette measure of cohesion and separation. Since the indicators HP and OP were continuous variables, we had the choice between log likelihood and Euclidean distance measures. In this article, we report the results based on Euclidean distances. As sensitivity analysis, we checked whether the results and conclusions remained the same with the use of (a) log likelihood distances in the two-step cluster analyses, and (b) Latent Profile Analyses (LPA). LPA were estimated using Mplus (version 7.1; Muthén & Muthén, 1998-2012), with log likelihood distance measures, and the robust MLR estimator. In the LPA, HP and OP were allowed to correlate, with equal co-variances assumed across clusters. LPA Models with two and more clusters were estimated and compared in terms of

AIC, BIC, entropy, Vuong-Lo-Mendell-Rubin Test, the Lo-Mendell-Rubin Adjusted LRT Test, and the bootstrap likelihood ratio difference test.

The clusters (resulting from the two-step cluster analysis with Euclidean distances) were interpreted based on their mean scores on the indicator variables HP and OP, and their mean scores in other validation criteria that are expected to be highly correlated with passion (commitment, the passion score of the com.pass scale, dispositional approach motivation/BAS, positive and negative affect, and symptoms of dependency). Differences between clusters on the mean scores of these variables were tested with *t*-tests and effect size *d* (see Table 5).

Results

Descriptives for all measures are summarized in Table 2. The mean score of HP is higher than the mean score of OP in all four studies.

Research question 1: Do individuals who are labeled "obsessive" according to the z-score-based procedure display higher raw-scores in OP than HP, and vice versa?

As the lower row of Figure 2 shows, the *z*-score based grouping procedure from Mageau et al. (2009); Philippe et al. (2009) and Vallerand and Houlfort (2003) created similar groups as found in these previous studies. The non-passionate individuals displayed levels of HP and OP far below the sample averages. The 'harmonious' group was characterized by above-average levels of HP and below-average levels of OP, whereas the 'obsessive' group was characterized by above-average levels of OP and below-average levels of HP. There were no individuals with identical levels of standardized HP and OP (see Table 4).

As the upper row of Figure 2 shows, we found that the group labels 'mainly harmonious' and mainly obsessive' were misleading, because both the individuals in the 'mainly harmonious' group and the individuals in the 'mainly obsessive' group displayed higher HP mean raw scores than OP mean raw scores on the original response scale in sample one, three and four. In sample two, the 'mainly obsessive group' reported very similar levels of HP and OP (mean 5.1 for both HP and OP). In other words: The individuals which are called 'mainly obsessive' by the z-score-based procedure report on average higher HP than OP raw scores when answering to the original response scale.

---

[3] In the process of this cluster analysis, the indicators are z-standardized by default. We found that this leads to slightly different person-cluster assignments than conducting the cluster analysis with raw scores, because it makes a difference whether absolute or relative distances between observations are analyzed. While the conclusion remained the same in our data example, this is a different problem of z-standardization in person-oriented analyses, to be addressed in future studies.
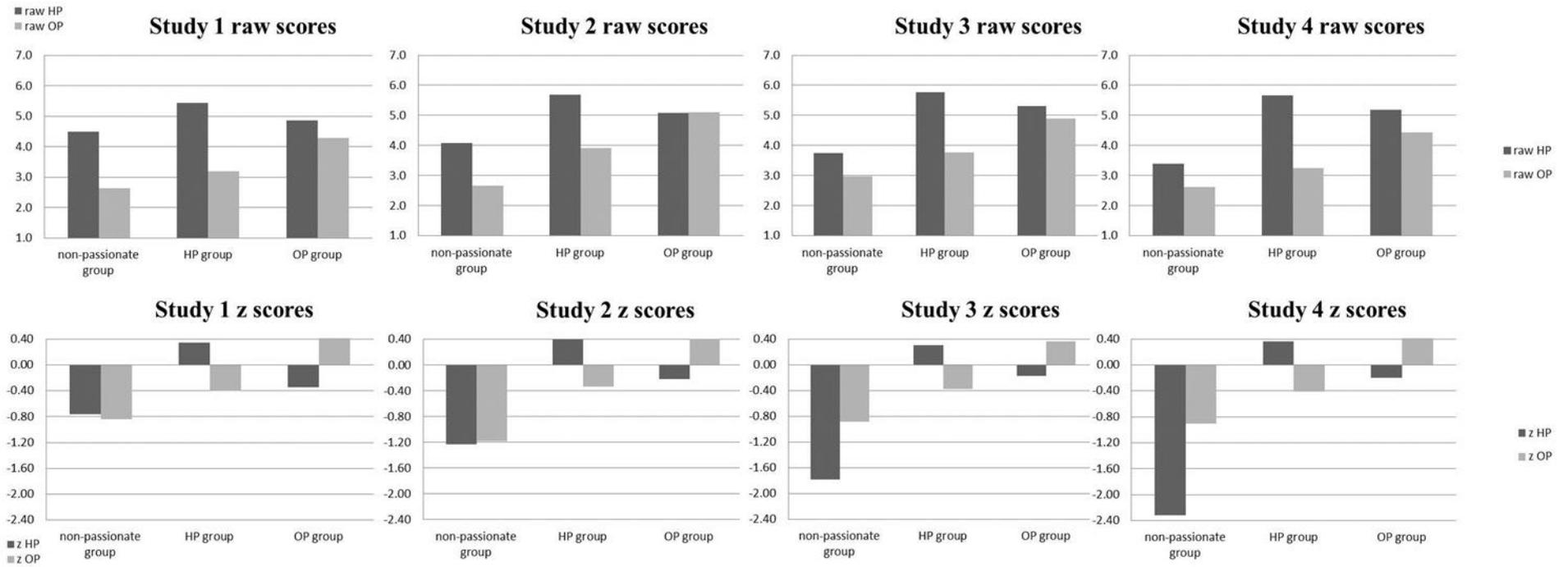
Figure 2: Mean raw scores (original response scale) for HP and HP per group and study (upper row), compared to mean z-scores of HP and OP per group and study (lower row).

Table 3

Percentages of misleadingly classified individuals

	Sample 1	Sample 2	Sample 3	Sample 4
$HP_{raw} > OP_{raw}$, but labelled 'obsessive' with the z-score procedure	38.8%	21.6%	32.6%	40.2%
$OP_{raw} > HP_{raw}$, but labelled 'harmonious' with the z-score procedure	0.0%	0.0%	0.0%	0.0%

Note. The percentages refer to the whole sample.

Next, Table 3 shows how many individuals with higher raw scores on HP than OP were classified as 'mainly obsessive' with the z-score- based procedure, and how many individuals with higher raw scores on OP than HP were classified as 'mainly harmonious' based on their z scores. Between 22% and 40% of all individuals were classified as 'mainly obsessive' with the z-score procedure despite of having reported higher HP than OP on the original response scale. The opposite misclassification of being classified as harmonious despite of stronger OP raw scores did not occur in any sample.

Research question 2: Do alternative classification procedures such as grouping based on raw scores and cluster analysis also produce the expected groups of mainly harmonious and mainly obsessive individuals?

Table 4 compares how many individuals are classified as non-passionate, harmonious, obsessive, and equally harmonious-obsessive with the *z*-score- based procedure versus the raw-score-based classification. While the *z*-score-based procedure suggest that each sample consisted of roughly as many harmonious as obsessive individuals (group sizes of 44.7% to 53.9%), the raw score based comparisons of HP and OP show that the vast majority of individuals (67.8% to 93.2%) reported higher HP than OP when answering to the original response scale.

In contrast to what we would expect based on the previous literature and based on the results of the above mentioned grouping procedures, the scatter plots indicated that it might not be too insightful to distinguish between different groups of individuals in terms of HP and OP, as both variables continually scattered around the regression line (see Figure 3). In most cases, HP and OP were either both low, or both moderate, or both high. HP generally tended to be higher than OP, in reference to the original response scale. Importantly, there were very few observations with low HP and high OP, as the empty right lower corners of the scatter plots indicate.

The two-step cluster analyses with Euclidean distances confirmed what the scatter plots had already indicated:

There were either no distinguishable groups (in samples 1 and 3), or two groups with one representing high passion, and the other representing low passion (in samples 2 and 4). Where a two-cluster-solution was found, the HP and OP levels within each cluster were aligned, meaning HP and OP were both relatively high in the 'high passion cluster', and both relatively low in the 'low passion cluster'. RawHP was higher than rawOP in each cluster. Discrepancy clusters with higher OP than HP were not found in any sample. The cluster solution was evaluated to be on the lower margin of "good", according to the silhouette measure of cohesion and separation (average silhouette = .5). That individuals display aligned scores of either high HP and OP, or moderate HP and OP, or low HP and OP was also supported in our sensitivity analyses of two-step cluster analysis with log likelihood distances (SPSS), and latent profile analyses (using log likelihood and the robust MLR estimator, and accounting for correlations between HP and OP, in Mplus). These alternative clustering procedures indicated again that there were no clearly pronounced profiles, but alignment of HP and OP scores, which suggests that the distinction between profiles was not more insightful than the concept of passion as a continuous phenomenon (for details see the appendix).

In the samples with a two-cluster solution, the two clusters differed in the general level of passion (generally high HP and OP versus generally low HP and OP), and these differences between clusters were significant and large according to the *t*-tests and effect sizes (see Table 5).

To validate the interpretation of these clusters, we examined how the clusters identified with the two-step cluster analyses in samples 2 and 4 differed in other constructs known to be related to passion, namely commitment, dispositional approach motivation, positive and negative affect and dependency. As expected, we found that the high passion cluster was characterized by high levels of passion, commitment, and dispositional approach motivation, while these constructs were significantly lower in the low passion cluster, the effect sizes were mostly large (see Table 5). The difference in regard to feeling dependent of the activity was not significant (sample 4).

Table 4

Percentages of non-passionate, harmonious (HP > OP), obsessive (OP > HP) and equally harmonious and obsessive (HP = OP) individuals: comparing the z-score-based classification with the raw score-based classification.

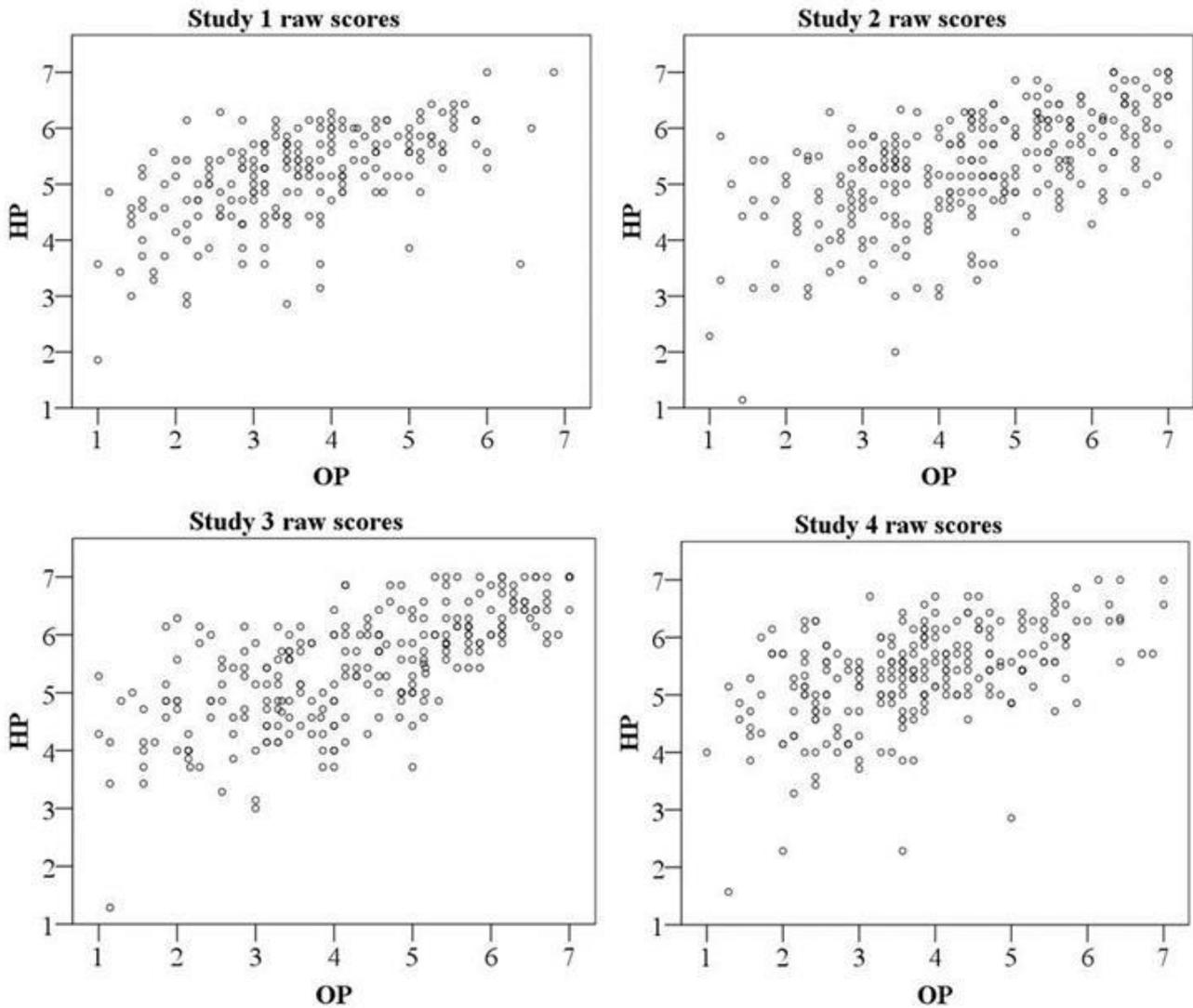|	Sample 1		Sample 2		Sample 3		Sample 4	
	z-score	Raw score	z-score	Raw score	z-score	Raw score	z-score	Raw score
Nonpassionate	1.5%	1.5%	5.9%	5.9%	2.7%	2.7%	3.7%	3.7%
HP > OP	53.9%	93.2%	46.2%	67.8%	44.7%	77.3%	49.2%	89.3%
OP > HP	44.7%	5.3%	48.0%	20.1%	52.7%	9.5%	47.1%	5.3%
HP = OP	0%	0%	0%	6.2%	0%	10.6%	0%	1.6%

Figure 3: Scatter plots of the raw scores of HP and OP in the samples 1 to 4.

Table 5

Comparing the high and low passion cluster in samples 2 and 4 (two-step cluster analyses, Euclidean distances)

	High Passion Cluster				Low Passion Cluster				t-tests of differences between the high and low passion cluster			
	Sample 2 N = 228		Sample 4 N = 233		Sample 2 N = 45		Sample 4 N = 11		Sample 2		Sample 4	
	M	SD	M	SD	M	SD	M	SD	p	d	p	d
Indicator HP [a]	5.61	0.674	5.46	0.694	3.66	0.709	3.16	0.799	.000	2.13	.000	1.37
Indicator OP [a]	4.72	1.358	3.85	1.267	2.85	0.980	2.49	1.135	.000	2.42	.001	0.45
General Passion criteria [a]	5.92	0.788	6.12	0.759	4.84	0.930	3.60	1.956	.000	0.99	.002	2.67
Passion [b]	4.78	0.663	4.53	0.642	3.88	0.685	3.62	0.586	.000	1.01	.000	0.53
Commitment	4.35	0.533			3.69	0.592			.000	0.90		
Dispositional approach motivation			3.11	0.369			2.77	0.394			.011	0.35
Dependency			3.16	0.758			2.74	0.816			.154	0.20

Note. This table reports differences between the two clusters which were found in the studies 2 and 4. In studies 1 and 3, no clusters were found. The t-tests require replication, due to the small sample size of the second cluster, particularly in study 4. [a] = Dual Model Passion Scale (Vallerand et al., 2003), [b] = com.pass scale (Moeller & Grassinger, 2014b).

Discussion

This study had two purposes: First, we investigated whether the two 'types of passion' describe distinct subgroups of individuals, or rather two aspects that, while explaining different outcomes, are experienced together by an individual. Second, we discussed why and when *z*-standardization might be misleading in the analysis of subgroups.

We examined the z-score-based grouping procedure that was used to distinguish mainly harmonious from mainly obsessive individuals in previous studies (Mageau et al., 2009; Philippe et al., 2009; Vallerand & Houlfort, 2003). With this procedure we found similar groups as in previous studies. However, we found that the interpretation of these groups was misleading, because many individuals were classified as 'obsessive' despite higher HP than OP in their original answers, which was due to the different mean scores of HP and OP. On average, the individuals of each sample reported much higher HP than OP. Z-standardization conceals these mean level differences, because the standardized mean score is always 0. The groups created with the z-score-based procedure do not reflect the individual's response reference frame for 'high' versus 'low' passion, but the individual's rank in reference to other individuals in the same sample. This implies that an individual with a high OP z-score might have experienced higher OP than other individuals, but probably still lower OP than HP.

We then applied alternative classification procedures to identify groups of individuals, namely the classification into a harmonious and an obsessive group based of raw scores of HP and OP, and cluster analyses. These analyses and the scatter plots revealed that (1) HP and OP are aligned for most individuals, meaning either both relatively high, or both relatively low, (2) most individuals report higher levels of HP than OP in terms of the original response scale, and (3) prototypical obsessive individuals with stronger OP than HP were rare in the raw-score-based grouping procedure, and not visible in the cluster analyses. These results were replicated in all four samples, despite the fact that the participants differed between the samples in regard to their passionate activity context, home country, and age.

We conclude that, in samples like ours of individuals who pursue activities on amateur level, the only useful distinction of subgroups is the distinction between generally highly passionate individuals, and generally lowly passionate individuals. It seems more appropriate to conceptualize passion as a continuous variable, with beneficial and harmful aspects that occur together. In sum, our findings suggest that passion is a double-edged sword. Individuals who report strong passion seem to experience at the same time harmonious and obsessive aspects, positive and negative affect, and voluntary commitment (see Tables 5 and 6). While the difference in dependency between the high and the low passion cluster was not significant in this study, this finding requires replication due to the small size of the second cluster in sample 4. Although obsessive experiences seem rare and not intense in individuals with low passion, they are more strongly pronounced in individuals with a strong passion. Nevertheless, the harmonious experiences prevail over the obsessive even for individuals with strong passions. The often assumed prototypical obsessed individuals with high OP and low HP (e.g. Vallerand et al., 2003; 2012) were not found.

Harmonious and obsessive experiences as two sides of the same coin

The dual model emphasizes the duality between HP and OP, while we found that most individuals experience either

high HP *and* high OP, or low HP *and* low OP. At first glance, the co-occurrence of HP and OP seems an oxymoron, as HP describes opposite experiences to OP (HP: maintained control, positive emotions, well-being; versus OP: loss of control, negative emotions, ill-being). However, HP and OP are correlated with each other, with effect sizes ranging from small (Marsh et al., 2013) to moderate (Rip et al., 2006; Vallerand et al., 2007) to high (Séguin-Lévesque et al., 2003; Vallerand et al., 2003; Vallerand, Ntoumanis, et al., 2008, study 1). This indicates that HP and OP co-occur for at least some individuals in many previous studies. To understand that, it helps to reflect on the fact that the ratings of passion are based on the individual's repeated experiences of an activity over time. Some individuals might experience alternatingly intrinsic and aversive states while engaging in their passionate activity. Likewise, individuals might experience autonomous and constraint reasons to identify with an activity together or alternatingly, so that the autonomous reasons might then lead to harmonious experiences, and the controlled reasons to obsessive experiences, within the same individual. Also, reasons to engage in an activity might develop and change over time, so that harmonious passions might become more obsessive or vice versa over time.

An important implication is that our findings do not support that a high level of HP was a generally desirable experience, as it often comes accompanied by high levels of OP. As Wang and Yang (2007, p. 297) put it: "high OP exists only when high HP occurs. This means that HP is a requirement of OP and HP and OP can be denoted as a sequence of different degrees of passion." It would be interesting to shed more light on the differences between highly and lowly passionate individuals, as these were rarely studied before. The scatter plots (see Figure 3) and the fact that HP and OP are aligned within clusters indicate that it might not be very insightful at all to distinguish between separate subgroups of individuals in terms of HP and OP.

Our findings indicate that it can be misleading to claim that individuals with high OP were at risk, while individuals with high HP were well adapted, since HP is often at least as high, often higher, than the OP in any given person. Interestingly, there is a 'positive bias' in all measures of passion: Most individuals affirm the items for HP much stronger than the items for OP (see Mageau et al., 2009; and above). Furthermore, OP was found to correlate with positive affect and other positive experiences (e.g., Moeller et al., 2014). Moreover, the presumably neutral scale of general passion criteria of the dual model passion scale correlates moderately to strongly with positive affect (Moeller et al., 2014). In sum, these findings suggest that passion is generally a predominantly positive experience, despite the obsessed feelings that often accompany a strong passion, and disregarding the different correlates of OP and HP.

Nevertheless, our findings corroborate the assumption that high motivation can have downsides, such as obsessive, risky and ill-advised persistence, and the risk of exhaustion.

Similar downsides of high motivation were described in recent studies on the relation of student engagement to student burnout, which showed that for a substantial subgroup of individuals, strong engagement comes accompanied by strong exhaustion and burnout symptoms (Tuominen-Soini & Salmela-Aro, 2014; Salmela-Aro et al., 2015), despite the fact that engagement has been described as optimal motivation and 'flourishing' experience in the previous research (e.g. Reschly, Huebner, Appleton, & Antaramian, 2008). The dual model of passion has helped us to perceive this "dark side of motivation", and future studies should consider the up- and downsides of motivational constructs in more detail.

It might be the context that makes the difference: Environments where individuals face much pressure and constraints, such as highly competitive sports, might produce or attract more mainly obsessive individuals than leisure sport teams where people drop out if they feel that their hobby interferes with their autonomy and well-being (Amiot, Vallerand & Blanchard, 2006). Possibly, highly competitive Olympic athletes might experience more obsession than harmonious passion in critical periods of intensive training. Future studies could shed light on the relation between contextual characteristics and the development of passion profiles. In any case, there are domains and samples where the assumed universal duality of passion in terms of pronounced profiles was not supported. Moreover, a recent study on young elite soccer players on the edge of entering the professional level found that even in such highly competitive situations, HP tends to be substantially higher than OP, across all individuals (Chamorro, Torregrosa, Sánchez-Oliva, León, & García Calvo, 2015).

Implications for the research on group differences

One implication of our study is that z-scores can be misleading in the analyses of profiles and groups. Z-standardization is often applied to compare the level of individual responses in different variables. While this is unproblematic in cross-sectional inter-individual correlation analyses, it becomes problematic for instance when variables with different distributions are first standardized, and then these standardized scores are used for the interpretation of different profiles. A graph of profiles should show the information about (1) the absolute item expression: which group affirmed or rejected which variable on average (which variable mean score was above/below the scale midpoint on a likert rating scale); (2) the relative item expression: which variable score was higher in which profile/group, and how large were these profile differences; and (3) whether the same profiles were found in different samples. Z-scores obfuscate all of these pieces of information that usually are read from a graph of profiles because of the following reasons:

(1) Sometimes z-scores suggest that response A was higher than response B, even though it was the other way around when the person answered to the original response scale. That implies that a "high" score can turn into a "low" score and vice versa, when raw scores are plotted instead of z-scores.

(2) *z*-scores can sometimes suggest that an answer was "high", even though this answer negated the item statement on the original response scale (= was a rating below the scale midpoint on a scale from 1 = *don't agree at all* to 7 = *totally agree*).

(3) Plotting profile differences using z-scores often equals plotting a graph with a truncated Y-axis, because rarely the whole area of possible answers is shown in such *z*-score graphs.

(4) *z*-scores are very sample-specific. The same z-score can represent widely different raw scores (i.e., different degrees of item affirmation) in different samples. If the sample is representative for the population, this might not be a problem, but in most studies samples are not representative in all relevant aspects, and then *z*-scores are not comparable across samples.

In other words, displaying z-scores instead of raw scores in profiles makes the central information of the plot unreliable. We suggest to report raw scores in addition to or instead of z-scores in analyses of differences between profiles or groups, because raw scores reflect the original frame of reference and an individual's subjective acceptance versus rejection of items much better than z-scores do, and raw scores provide the information about absolute distances between two measurements, whereas z-scores obfuscate this information (see also Moeller, 2015). Interpreting group or profile differences based on raw scores requires that the scales used for comparison are indeed comparable. Reasons why they might not be comparable can be (1) the items were measured with different response scales and metrics (e.g. scale A ranged from 1 = *don't agree at all* to 7 = *totally agree* and scale B ranged from 0 = *don't agree at all* to 10 = *totally agree*.), or (2) Even when the response scales were the same, the item wording of one scale might be more extreme than another, e.g. scale A might ask for moderate symptoms of depression while scale B might ask for severe symptoms of depression. In that case, a high score on scale A (moderate depression symptoms) would not describe the same severity of depression as an equally high score on scale B (severe symptoms).

For the first issue (different response scales) there is an easy solution: If variables with different metrics are used, they can be brought to the same metric with alternative transformations that reflect the original response scale better, as for example the proportion of maximum scaling (POMS; Little, 2013) which makes each scale range from 0 to 1 while maintaining the proportions of absolute distances (for an SPSS syntax, see Moeller, 2015). Since the POMS transformation preserves the information about absolute distances, it might be that this procedure is more useful than *z*-standardization not only for the interpretation of clusters, but already for the estimation of clusters.

Regarding the second issue (differences in extremeness of item formulation), it is useful to compare the item difficulties of the applied measures. When scale B generally has a lower mean score than scale A on a given response scale metric, then this indicates that the items of scale B are generally more difficult, that is, less likely to be affirmed than items of scale A. The question is then: Do individuals affirm the items of scale B less than the items of scale A because they really do not experience the symptoms described by scale B, or do individuals affirm scale B less than scale A, because scale B uses more extreme and drastic descriptions of the same symptoms that scale A describes (e.g., scale B: "I feel sad all the time", scale A: "I sometimes feel sad")? One possibility to answer this tricky question is to carefully compare the item wording of the applied scales in terms of differences in item extremeness. Another possibility is to apply item response theory (IRT) to get independent estimates of the item difficulty and the latent trait of the person. This article is based on the assumption that the mean scores of HP and OP were comparable, because (1) both scales were rated on the same response scale and the items were presented in a mixed order, and (2) the item wordings of HP and OP are comparable in terms of extremeness. The only words that make OP items sound more extreme than HP items describe aspects that are at the core of the definition of obsessive passion ('urge', 'almost obsessive feeling', preferring this activity over all others, 'sometimes lose control', 'activity controls me'). Thus, the different item difficulties of HP and OP can be interpreted as direct reflections of different experiences of the constructs as they are defined.

It is crucial, and a question of face validity, whether two scales differ in mean scores due to words that are related to the core of the construct's definition, or due to words that are unrelated to the core definition of the construct and rather describe a severity. In the latter case, even raw or POMS-transformed scores would not necessarily be comparable. One possible solution is to first make sure that both scales measure the exact same construct on different levels of extremeness through validity analyses, and then weighting a person's raw scores of each scale at the interpersonal item difficulty of this scale (or its reciprocal value, respectively) in order to put the original rating into perspective. (The difficulty of an item, or scale, can be computed using almost the same formula as POMS transformation applies)[4].

---

[4] POMS = [(*observed* - *minimum*)/(*maximum* - *minimum*)]; Item difficulty = [(*MEAN* - *minimum*)/(*maximum* - *minimum*)]*100.

Limitations

Since our measures were based on self-report questionnaires, we cannot rule out that differences between individuals might be blurred or amplified by response styles, such as a general tendency to affirm any item disregarding the specific content. Such acquiescence might explain the tendency to respond in the same way to harmonious and obsessive items, which in turn could explain the clusters of aligned high or low passion scores. However, this would apply to any self-report measure, and concerns most studies in the previous research on passion as well, since we do not know of passion studies dealing with the acquiescence problem. Therefore, we would argue that the findings of this study provide important insights to the passion research as we know it. Methods to reduce response-style biased findings could be control variables or method factors, for instance in factor-mixture models (see Billiet & McClendon, 2000; Geiser & Lockhart, 2012).

The relatively small sample sizes of our samples limit the trustworthiness of the cluster analyses. However, five arguments support our conclusions: (1) all four cluster analyses pointed to the same result; (2) the same results are found if our four samples are merged into larger samples; (3) we found the same clusters of aligned high passion versus aligned low passion and no obsessive group in two much larger samples in Finland (each more than 1,000 individuals; Moeller, Hietajärvi, Tuominen-Soini, Lonka, & Salmela- Aro, K., 2015); (4) our findings replicate earlier results from Wang et al. (2008), who found HP and OP aligned in their clusters (see also Wang & Yang, 2007); and (5) our conclusions remained the same with different clustering procedures. Thus, our findings seem generalizable across different cluster methods, activity contexts, age groups and countries/language versions of the dual model passion scale.

Authors' Contributions

JM had the leading role in planning and conducting all four studies and data analysis for this article and writing the manuscript. MK collaborated in the planning, preparation, data collection and analysis of samples three and four, and carried out robustness checks and helped revising the manuscript. RG contributed to the data collection in sample one and to the development of the com.pass scale, used as validation instrument in all samples.

Acknowledgments

References

Amiot, C. E., Vallerand, R.J., & Blanchard, C. M. (2006). Passion and psychological adjustment: A test of the person-environment fit hypothesis. *Personality and Social Psychology Bulletin*, 32 (2), 220-229. doi:10.1177/0146167205280250

Bergman, L. R., & Magnusson, D. (1997). A person- oriented approach in research on developmental psychopathology. *Development and Psychopathology, 9(2)*, 291–319.

Billiet J. B. & M. J. McClendon (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling: A multidisciplinary Journal, 7(4)*, 608–628. doi:10.1207/S15328007SEM0704_5

Cardon, M. S., Wincent, J., Singh, J., & Drnovsek, M. (2009). The nature and experience of entrepreneurial passion. *Academy of Management Review, 34(3)*, 511–532. doi:10.5465/AMR.2009.40633190

Carver, C. S. & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67(2)*, 319-333. doi:10.1037/0022-3514.67.2.319

Chamorro, J. L., Torregrosa, M., Sánchez-Oliva, D., León, B. & García Calvo, T. (2015). Future achievements, passion and motivation in the transition from junior-to-senior sport in Spanish young elite soccer players. Manuscript submitted for publication.

Fredricks, J. A., Alfeld, C., & Eccles, J. S. (2010). Developing and fostering passions in academic and nonacademic domains. *Gifted Child Quarterly, 54(1)*, 18-30. doi:10.1177/0016986209352683

Frijda, N. H. (2007). *The Laws of Emotion.* Mahwah, NJ: Lawrence Erlbaum Associates.

Geiser, C., & Lockhart, G. (2012). A Comparison of Four Approaches to Account for Method Effects in Latent State-Trait Analyses. *Psychological Methods, 17(2)*, 255–283. doi:10.1037/a0026977

Kessler, E.-M., & Staudinger, U. M. (2009). Affective Experience in Adulthood and Old Age: The Role of Affective Arousal and

Perceived Affect Regulation. *Psychology and Aging, 24(2)*, 349-362. doi:10.1037/a0015352

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchungen mit einer deutschen Version der „Positive and Negative Affect Schedule" (PANAS) [Investigations with a German version of the Positive and Negative Affect Schedule (PANAS)]. *Diagnostica, 42(2)*, 139-156.

Little, T. D. (2013). *Longitudinal Structural Equation Modeling (Methodology in the Social Sciences)*. The Guilford Press.

Mageau, G. A., Vallerand, R. J., Charest, J., Salvy, S.-J., Lacaille, N., Bouffard, T., et al. (2009). On the development of harmonious and obsessive passion: The role of autonomy support, activity specialization, and identification with the activity. *Journal of Personality, 77(3)*, 601-646. doi:10.1111/j.1467-6494.2009.00559.x

Mageau, G., Vallerand, R., Rousseau, F., Ratelle, C., & Provencher, P. (2005). Passion and gambling: Investigating the divergent affective and cognitive consequences of gambling. *Journal of Applied Social Psychology, 35(1)*, 100-118. doi:10.1111/j.1559-1816.2005.tb02095.x

Marsh, H. W., Vallerand, R. J., Lafreniere, M. A. K., Parker, P., Morin, A. J. S., Carbonneau, N., et al. (2013). Passion: Does One Scale Fit All? Construct Validity of Two-factor Passion Scale and Psychometric Invariance over Different Activities and Languages. *Psychological Assessment,* 25(3), 796-809. doi: 10.1037/a0032573.

Millon, T., Davis, R., Millon, C., Wenger, A., Van Zuilen, M., Fuchs, M., & Millon, R. (1996). *Disorders of personality: DSM IV and beyond*. New York: Wiley.

Moeller, J. (2015). A word on standardization in longitudinal studies: Don't. *Frontiers of Psychology, 6* (1389). doi: 10.3389/fpsyg.2015.01389.

Moeller, J. (2015). *SPSS-Syntax for POMS- and POMP-transformation*. Retrieved from osf.io/eaxgb

Moeller, J., & Grassinger, R. (2014a). A Review of Passion Concepts and their Overlaps with Commitment: Opening a can of worms. In: J. Moeller (2014). Passion as concept of the psychology of motivation. Conceptualization, assessment, inter-individual variability and long-term stability (pp. 13-52). *Dissertation published online at http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-29036/DissJuliaMoeller.pdf*

Moeller, J., & Grassinger, R. (2014b): Measuring passion and commitment with one joint scale: psychometric properties and validity of the com.pass scale. In: J. Moeller (2014). Passion as concept of the psychology of motivation. Conceptualization, assessment, inter-individual variability and long-term stability (pp. 53-110). *Dissertation published online at http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-29036/DissJuliaMoeller.pdf*

Moeller, J., Keiner, M., & Wächter, D. (2014). A disposition for passion? Linking passionate approach motivation to facets of extraversion and non-substance related dependency. In: J. Moeller (2014). Passion as concept of the psychology of motivation. Conceptualization, assessment, inter-individual variability and long-term stability (pp. 111-157). *Dissertation pub-

lished online at http://www.db-thueringen.de/servlets/DerivateServlet/Derivate-29036/DissJuliaMoeller.pdf*

Moeller, J., Hietajärvi, L., Tuominen-Soini, H., Lonka, K., & Salmela-Aro, K. (2015). Not all sunshine and rainbows: Why even a harmonious passion is risky for well-being. Manuscript in preparation.

Molenaar, P. C. M. (2004). A Manifesto on Psychology as Idiographic Science: Bringing the Person Back Into Scientific Psychology, This Time Forever. *Measurement: Interdisciplinary Research and Perspectives, 2(4)*, 201-218. doi:10.1207/s15366359mea0204_1

Molenaar, P. C. M., & Campbell, C. G. (2009). The New Person-Specific Paradigm in Psychology. Current directions in Psychological Science, 18(2), 112-117. doi:10.1111/j.1467-8721.2009.01619.x

Muthén, L.K., & Muthén, B.O. (1998-2012). *Mplus User's Guide. Seventh Edition*. Los Angeles, CA: Muthén & Muthén.

Philippe, F., & Vallerand, R. J. (2007). Prevalence rates of gambling problems in montreal, canada: A look at old adults and the role of passion. *Journal of Gambling Studies, 23(3)*, 275-283. doi:10.1007/s10899-006-9038-0.

Philippe, F. L., Vallerand, R.J., Houlfort, N., Lavigne, G. L., & Donahue, E. G. (2010). Passion for an activity and quality of interpersonal relationships: The mediating role of emotions. *Journal of Personality and Social Psychology, 98(6)*, 917-932. doi:10.1037/a0018017

Philippe, F., Vallerand, R.J., & Lavigne, G. (2009). Passion does make a difference in people's lives: A look at well-being in passionate and non-passionate individuals. *Applied Psychology: Health and Well-Being, 1*, 3-22. doi:10.1111/j.1758-0854.2008.01003.x

Renzulli, J. S., Koehler, J. L., & Fogarty, E. A. (2006). Operation Houndstooth intervention theory: Social capital in today's schools. *Gifted Child Today, 29(1)*, 14-24.

Reschly, A. L., Huebner, E. S., Appleton, J. J., & Antaramian, S. (2008). Engagement as flourishing: The contribution of positive emotions and coping to adolescents' engagement at school and with learning. *Psychology in the Schools, 45(5)*, 419-431. doi:10.1002/pits.20306

Rip, B., Fortin, S., & Vallerand, R. (2006). The relationship between passion and injury in dance students. *Journal of Dance Medicine & Science, 10(1,2)*, 14-20.

Reizle, M. (2013). Introduction: Doubts and insights concerning variable- and person-oriented approaches to human development. *European Journal of Developmental Psychology, 10(1)*, 1-8. doi:10.1080/17405629.2012.742848

Salmela-Aro, K., Moeller, J., Schneider, B.; Spicer, J., & Lavonen, J. (2015). Integrating the light and dark sides of student engagement with person-oriented and situation-specific approaches. Manuscript submitted for publication.

Scanlan, T. K., Russell, D. G., Wilson, N. C., & Scanlan, L. A. (2003). Project on elite athlete commitment (PEAK): I. Introduction and Methodology. *Journal of Sport and Exercise Psychology, 25(3),* 360-376.

Séguin-Lévesque, C., Laliberté, M.L., Pelletier, L.G., Vallerand, R.J., & Blanchard, C. (2003). Harmonious and obsessive passions for the Internet: Their associations with couples' relationships. *Journal of Applied Social Psychology*, *33(1)*, 197-221. doi:10.1111/j.1559-1816.2003.tb02079.x

Simourd, D.J., Hoge, R.D., Andrews, D.A., & Leschied, A.W. (1994). An empirically-based typology of male young offenders. *Canadian Journal of Criminology, 36(4)*, 447-461.

Solomon, R. C. (1993). *The Passions. Emotions and the Meaning of Live.* Indianapolis, Indiana: Hacket Publishing Company.

Sousa, C., Viladrich, C., Gouveia M. J., Torregrosa M., & Cruz J. (2008). Estrutura interna da versão portuguesa do Sport Commitment Questionnaire (SCQ) e Funcionamento Diferencial dos Itens (DIF) com a versão espanhola. [The internal structure of the Portuguese version of the Sport commitment questionnaire (SCQ) and differential item functioning (DIF) of the items of the Spanish version]. *Revista Portuguesa de Ciências do Desporto [online], 8(1)*, 22-33. Retrieved from: http://www.scielo.gpeari.mctes.pt/scielo.php?script=sci_arttext&pid=S1645-05232008000100003&lng=pt.

Sternberg, R. J. (1997). Construct validation of a triangular love scale. *European Journal of Social Psychology, 27(3)* , 313-335. doi:10.1002/(SICI)1099-0992(199705)27:3<313::AID-EJSP824>3.0.CO;2-4

Strobel, A., Beaducel, A., Debener, S., & Brocke, B. (2001). Psychometrische und strukturelle Merkmale einer deutschsprachigen Version des BIS/BAS-Fragebogens von Carver und White [Psychometric and structural properties of a German version of the BIS/BAS scale from Carver and White]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 22*, 216-227.

Tuominen-Soini, H., & Salmela-Aro, K. (2014). Schoolwork engagement and burnout among Finnish high school students and young adults: Profiles, progressions and educational outcomes. *Developmental Psychology, 50*(3)*, 649–662. doi:10.1037/a0033898

Terry, A., Szabo, A., & Griffiths, M. (2004). The Exercise Addiction Inventory: A new brief screening Tool. *Addiction Research and Theory, 12 (5)*, 489-499. doi:10.1080/16066350310001637363

Vallerand, R. J. (2012). The role of passion in sustainable psychological well-being. *Psychology of Well-Being: Theory, Research and Practice*, 2(1). doi:10.1186/2211-1522-2-1

Vallerand, R. J., Salvy, S.-J., Mageau, G. A., Elliot, A. J., Denis, P. L., Grouzet, F. M., et al. (2007). On the role of passion in performance. *Journal of Personality, 75(3)* , 505-534. doi:10.1111/j.1467-6494.2007.00447.x

Vallerand, R. J., Mageau, G. A., Elliot, A. J., Dumais, A., Demers, M.-A., & Rousseau, F. (2008). Passion and performance attainment in sport. *Psychology of Sport and Exercise, 9(3)*, 373–392. doi:10.1016/j.psychsport.2007.05.003

Vallerand, R.J., Ntoumanis, N., Philippe, FL.., Lavigne, G.L., Carbonneau, N., Bonneville, A., et al. (2008). On passion and sports fans: A look at football. *Journal of Sport Sciences. 26(12)*, 1279-1293. doi:10.1080/02640410802123185

Vallerand, R.J., & Houlfort, N. (2003). Passion at work: Toward a new conceptualization. In D. Skarlicki, S. Gilliland, & D. Steiner (Eds.), *Emerging Perspectives on Values in Organizations (pp. 175–204).* Greenwich, CT: Information Age Publishing.

Vallerand, R. J., Blanchard, C., Mageau, G. A., Koestner, R., Ratelle, C., Léonard, M., et al. (2003). Les passions de l'âme: On obsessive and harmonious passion. *Journal of Personality and Social Psychology, 85(4)*, 756-767. doi:10.1037/0022-3514.85.4.756

Von Eye, A., & Bergman, L. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology, 15(3)*, 553–580. doi:10.1017.S0954579403000294

Wang, C. K. J., Khoo, A., Liu, W. C., & Divaharan, S. (2008). Passion and intrinsic motivation in digital gaming. *Cyberpsychology and Behavior, 11(1)*, 39-45. doi:10.1089/cpb.2007.0004

Wang, C.C., & Yang, H.W. (2007). Passion and dependency in online shopping activities. *Cyberpsychology and Behavior 10(2)*, 296–298. http://dx.doi.org/10.1089/cpb.2006.9954.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54(6)*, 1063-1070. doi:10.1037/0022-3514.54.6.1063

Appendix: Results of the Latent Profile Analyses

Summary: The appendix reports the results of the latent profile analyses, which were run to check the replicability of our findings. Table 6 summarizes the fit indices, and Figure 4 displays the mean scores of the cluster solutions which were identified as best descriptions of our data. In sum, the latent profile analyses describe no pronounced profiles, but rather alignment of HP and OP scores, which replicates our previous findings. The number of identified clusters varies across samples, but the conclusion remains the same.

In sample 1, the model with 2 clusters already did not describe the data better than an assumed 0-cluster model (according to VLMR, VLMR adj., and Bootstrap likelihood test), ergo, no distinction between subgroups was insightful. In sample 2, the VLMR, VLMR adj., and Bootstrap likelihood test support the 2- and 3-cluster solutions, the entropy the 2-cluster solution, and the AIC the 3- or 4-cluster solution. In sample 3, the VLMR and VLMR adj. speak in favor of the 2-cluster solution, whereas the Bootstrap likelihood test supports the 2-, 3-, and 4-cluster solutions, the AIC the 5-cluster solution and the BIC the 2-cluster solution. In study 4, the VLMR, VLMR adj., and Bootstrap likelihood test speak for the 2-cluster solution.

s

Table 6
Model summaries for the latent profile analyses

| Sample	Number of clusters	Parameters	Log L	Corr. factor	AIC	BIC	Entropy	VLMR	VLMR adj.	Bootstrap likelihood test	Average Latent Class Probabilities for Most Likely Latent Class Membership by Latent Class	Smallest cluster
1	2	8	-537.75	1.2251	1091.501	1118.124	0.871	0.0786	0.0878	0.0000	0.853	9.7%
	3	11	-533.745	1.1298	1089.490	1126.097	0.781	0.2716	0.2842	0.3750	0.757	4%
2	2	8	-806.961	1.0271	1629.921	1658.797	0.813	0.0157	0.0190	0.0000	0.853	9.9%
	3	11	-798.54	1.0535	1619.080	1658.784	0.688	0.0147	0.0190	0.0000	0.831	9.5%
	4	14	-795.169	0.8905	1618.338	1668.87	0.818	0.1110	0.1207	0.5000	0.863	0.7%
3	2	8	-760.944	1.0705	1537.889	1566.587	0.691	0.0072	0.0093	0.0000	0.872	33.3%
	3	11	-752.715	1.2991	1527.430	1566.890	0.649	0.375	0.3921	0.0128	0.767	18.0%
	4	14	-746.567	0.9466	1521.133	1571.355	0.745	0.0639	0.0694	0.0200	0.815	0.4%
	5	17	-741.956	0.9067	1517.912	1578.896	0.724	0.1045	0.1164	0.0811	0.687	0.4%
4	2	8	-661.264	0.8679	1338.527	1366.505	0.989	0.0000	0.0000	0.0000	0.997	1.6%
	3	11	-657.817	0.8657	1337.634	1376.103	0.734	0.0708	0.0837	0.4286	0.784	1.6%
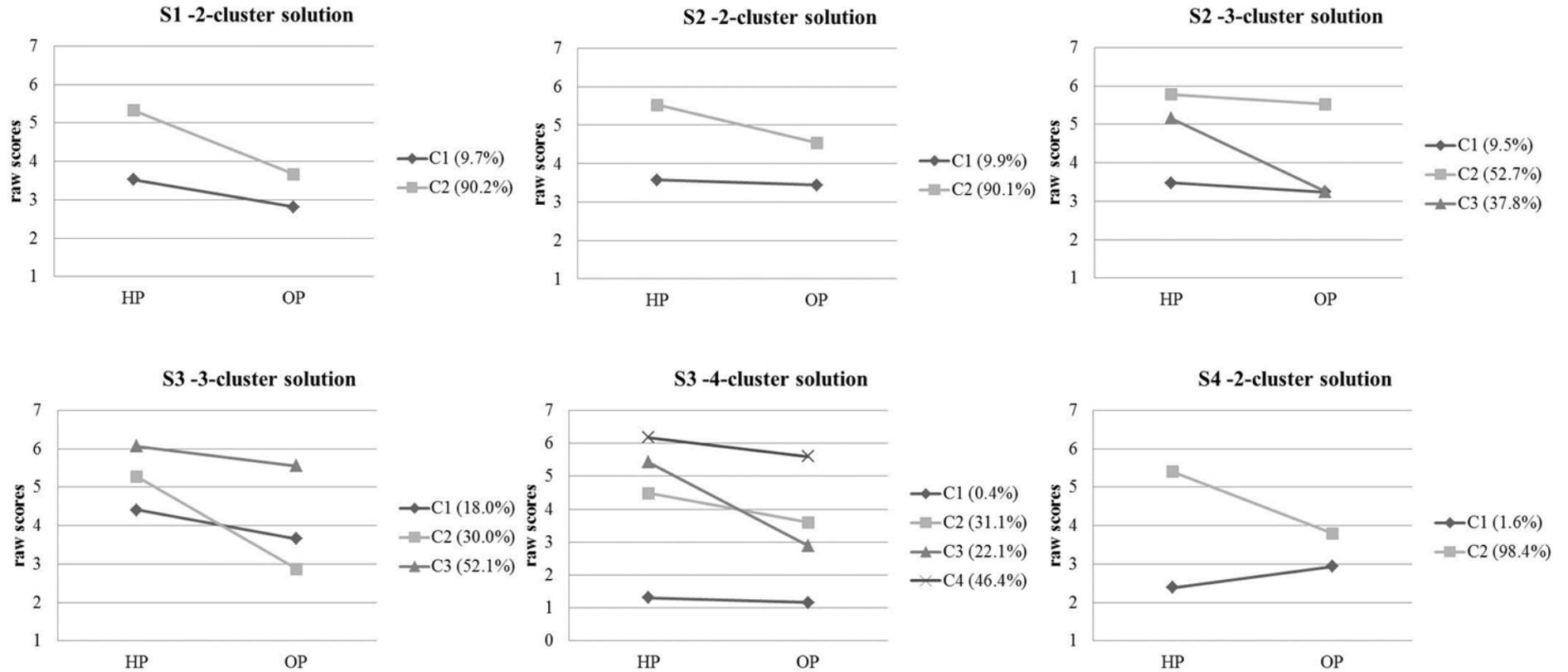
Figure 4: Mean scores of HP and OP per profile according to the LPA. Because of the unclear results regarding the number of clusters in samples 2 and 3, we display both plausible solutions for these studies.