Empathy trajectories throughout medical school: relationships with personality and motives for studying medicine

Piumatti, Giovanni; Abbiati, Milena; Baroffio Barbier, Anne; Gerbase, Margaret

How to cite

PIUMATTI, Giovanni et al. Empathy trajectories throughout medical school: relationships with personality and motives for studying medicine. In: Advances in Health Sciences Education, 2020. doi: 10.1007/s10459-020-09965-y

This publication URL: https://archive-ouverte.unige.ch//unige:137152
Publication DOI: 10.1007/s10459-020-09965-y

Empathy trajectories across medical studies: Relationships with personality and motives for studying medicine

Giovanni Piumatti,[1,2*] Milena Abbiati,[1] Anne Baroffio,[1] Margaret W. Gerbase[1]

[1] Faculty of Medicine, Unit of Development and Research in Medical Education (UDREM), University of Geneva, Geneva, Switzerland

[2] Population Epidemiology Unit, Primary Care Division, Geneva University Hospitals, Geneva, Switzerland

[*]Correspondence concerning this article should be addressed to Giovanni Piumatti, Unit of Development and Research in Medical Education (UDREM), Centre Médical Universitaire, 1, rue Michel Servet, CH-1211 Geneva 4, Switzerland, Tel.: +41 22 379 59 44, Email: giovanni.piumatti@unige.ch

Abstract

Empathy remains a widely discussed topic within medical education research. Studies on empathy changes among medical students are not univocal: empathy may decline, remain stable or increase. A largely unexplored research question regards inter-individual variability in empathy change, namely if different longitudinal trajectories of empathy exist. Evidence on the association of empathy trajectories with personality and motives for studying medicine is also scarce. Here, latent growth modeling examined empathy (measured with the Jefferson Scale of Empathy) among 201 medical students (M_{age}=20.74, 57% females) across three assessments: at entry year (Year 1) and during the first two clinical years (Years 4 and 5). Associations between empathy trajectories, personality in Year 1 and motives for studying medicine in Years 4 and 5 were tested. We identified two empathy trajectory groups: *lower and decreasing* (*n*=59; 29%) and *higher and stable* (*n*=142; 71%). Regression analyses indicated that higher openness in Year 1 was associated with an increased probability of *higher and stable* group membership (controlling for motives in Year 1). The effect of openness disappeared controlling for motives in Years 4 and 5 while caring for patients (in Years 4 and 5) and altruism (in Year 4) were positively associated with an increased probability of *higher and stable* group membership. In sum, we observed that empathy remains stable in most medical students and declines in fewer; openness and patients-oriented motives for studying medicine are associated with higher and stable empathy. Encouraging medical students' patients-oriented motives from preclinical throughout clinical years may prevent empathy decline.

Keywords: empathy, longitudinal trajectories, personality, motives, undergraduate medical students

Empathy trajectories throughout medical school: Relationships with personality and motives for studying medicine

Empathy remains a widely discussed topic and it has also been found to be a key correlate of medical students' career choices (Hojat et al., 2005; Tavakol et al., 2011), clinical competence (Chen et al., 2010; Hojat et al., 2002), emotional intelligence (Austin et al., 2005; Satterfield and Hughes, 2007), communication skills (Winefield and Chur-Hansen, 2000) and psychological distress (Brazeau et al., 2010; Dyrbye et al., 2006). Determining with certainty whether it is more likely to change or to remain stable throughout medical studies has proven to be inconclusive (Costa et al., 2013; Ferreira-Valente et al., 2017; Hegazi and Wilson, 2013; Hong et al., 2012; Neumann et al., 2011; Ponnamperuma et al., 2019; Quince et al., 2011; Spatoula et al., 2019; Wellbery et al., 2019). As a result, its trajectory course is still poorly understood, with possible decrements occurring during clinical training (Chen et al., 2012; Hojat et al., 2004; Papageorgiou et al., 2018; Ward et al., 2012). Moreover, the vast majority of previous longitudinal studies on this matter have adopted group-level analytical approaches that are unable to assess both inter- and intra-individual changes in empathy (Ferreira-Valente et al., 2017).

Inter-individual change or variability refers to differences between individuals, whereas intra-individual change refers to differences within the same person over time. While longitudinal and cross-sectional research has linked between-student differences to, among others, gender (Ferreira-Valente et al., 2017; Hojat et al., 2002; Neumann et al., 2011; Williams et al., 2014), personality (Costa et al., 2014; Hojat et al., 2005; Magalhães et al., 2012) and motives for studying medicine (Gonçalves-Pereira et al., 2013; Piumatti et al., 2018), to the best of the our knowledge, only one study by Costa et al. (2013) has examined in more detail issues of both between- and within-student longitudinal changes in empathy. In particular, although they pointed out non-significant within-individual changes (i.e. empathy

did not increase or decline over time) but significant inter-individual differences due to gender and personality, they did not explore further the significant observed variability in empathy growth rates. More specifically, their results suggested the possibility of different empathy trajectory groups within their population of medical students. Accordingly, testing whether different empathy trajectories may exist throughout medical studies can provide additional insights on empathy changes in relation to specific individual characteristics.

In the current study we examined between- and within-student changes in empathy during medical studies from the entry year to clinical training at the fourth and fifth academic years by further testing whether different trajectory groups could be observed. We then explored the determinants of trajectories, focusing on specific students' individual characteristics. In this regard, several studies showed a positive relationship between empathy and the personality traits of openness to experience and agreeableness within the Five-Factor Model (Costa et al., 2014; Costa et al., 2013; Costa and MacCrae, 1992; Guilera et al., 2019; Magalhães et al., 2012). Moreover, recent studies have pointed out associations between motivational factors and empathy levels among medical students: internal and altruistic motives for studying medicine (e.g. helping others, taking care of patients) rather than external and utilitarianism ones (e.g. economic return, social status) are associated with greater empathy (Gonçalves-Pereira et al., 2013). In particular, it has been observed that students showing a greater interest in patients-oriented medical specialties also report higher levels of empathy (Chen et al., 2007; Hojat et al., 2005; Vaglum et al., 1999) and that medical students' motives for studying medicine when entering medical school explain their self-reported empathy (Piumatti et al., 2018). In sum, we focused on personality in the entry year and on students' endorsement of different motives for choosing medicine across study years to explain their membership in empathy trajectory groups.

Method

Participants

The sample consisted of 201 medical school students from the Faculty of Medicine of the University of Geneva (Switzerland) recruited during two consecutive academic years in 2011 (Cohort 1, *n* = 100) and 2012 (Cohort 2, *n* = 101). For the current analyses, we selected participants who were admitted to the second academic year (selection took place at the end of Year 1, namely in 2012 for Cohort 1 and in 2013 for Cohort 2) and who took part in at least one of the following two data collections in their fourth (Year 4) and fifth (Year 5) academic years. As a result, our longitudinal sample included 62% of the total number of students who were successfully admitted to their second academic year between 2012 and 2013 (i.e. N = 324). The average age was 20.74 years (ranging from 18 to 38 years; SD = 2.21) and 57% (*n* = 114) were females. Participants filled in paper-and-pencil questionnaires in the interval between lectures. Before agreeing to take part in the study by signing a consent form, they were informed about the content of the research project, their entitlements and commitments as voluntary participants and the terms of confidentiality and privacy. Participants provided their student ID in order to be matched throughout the duration of the study. Researchers did not have simultaneous access to the data and student IDs, as the latter were managed by a technical administrator. The Chair of the Cantonal Commission for Ethical Research (CCER) designated the current study as exempted from formal review.

Educational context and empathy teaching

The undergraduate curriculum in medical studies lasts six years at the University of Geneva. From the entry year students are presented with courses dealing with the importance of the doctor-patient relationship from an interdisciplinary and communication-based perspective to explore the various dimensions of the clinical encounter: language, discourse and non-verbal communication. These aspects are further taught during the second and third

years by means of interactions with standardized and instructor patients across different themes and assessed by Objective Structured Clinical Examinations (OSCE). The fourth and fifth years consist of clinical clerkship rotations in different fields of medicine where students are taught by case-based clinical reasoning tutorials two days a week and take care of real patients under supervision when assigned to a ward unit three days a week. In the sixth and last year, undergraduate medical training concludes with a licensing exam.

Measures

Personality. The French version of the NEO Five Factor Inventory (NEO-FFI) (Aluja et al., 2005; Costa and MacCrae, 1992) was used to measure personality in Year 1. Five personality dimensions (each measured by 12 items on a 5-point Likert scale from 0 = *strongly disagree* to 4 = *strongly agree*) were measured: neuroticism (Cronbach's $\alpha = 0.87$), extraversion ($\alpha = 0.67$), openness ($\alpha = 0.73$), agreeableness ($\alpha = 0.78$) and conscientiousness ($\alpha = 0.85$). In contrast to motivation for studying medicine (Duffy et al., 2011; Gąsiorowski et al., 2015), personality scores are likely to remain stable across study years (Corker et al., 2013; Robins et al., 2001). Accordingly, for the current analyses, personality assessment was limited to Year 1.

Empathy. Empathy was measured in Years 1, 4 and 5 with the student's version of the Jefferson Scale of Empathy (JSE-S) (Hojat et al., 2001) translated into French (Zenasni et al., 2012). The JSE-S is composed of 20 items scored on a 7-point Likert scale (1 = *strongly disagree* to 7 = *strongly agree*). A total score for empathy was calculated by summing all answers to each question (α at Year 1 = 0.73; α at Year 4 = 0.83; α at Year 5 = 0.83).

Motives for studying medicine. In Years 1, 4 and 5, a list of different motives for studying medicine was presented to the students who were asked to indicate how important each of these were for themselves (i.e. "Describe how important each of these keywords is for your choice of medicine.") on a 6-point Likert scale (from 1 = *not important at all* to 6 =

very important): *reward*, *prestige*, *private practice*, *saving lives, caring for patients* and *altruism*. This set of motives was chosen based on a review of the research literature (Crossley and Mubarik, 2002; Goel et al., 2018; Lefevre et al., 2010; Williams and Deci, 1996), in order to have a wide enough description of different typologies of motives for choosing medicine. A similar approach was adopted by Vaglum et al. (1999) when studying motivation for attending medical school in a nationwide sample of medical students in Norway.

Data analyses

Preliminary analyses tested for accuracy of data entry, significant differences between cohorts, patterns of missing data and normality assumptions across assessments. Longitudinal analyses followed three steps. First, latent growth modeling (LGM) (Bollen and Curran, 2006; Duncan et al., 2013) estimated the average initial level and slope of empathy across Years 1, 4 and 5. Subsequently, analyses were extended into group-based trajectory modeling (GBTM) (Nagin, 2005) to evidence heterogeneity in empathy development. A complete explanation of this method and procedure in the current study is described in Appendix S1. Briefly, LGM examines trajectories of change in an outcome variable across time (empathy in this study). While LGM assumes that growth curves are representative of a single population, the underlying assumption when using GBTM is that the observed population can be classified into a finite number of different trajectory groups who exhibit similar patterns of change across time (Griffin et al., 2018). Thus, if LGM shows that on average students tend to remain stable in empathy across time, applying GBTM can further tell us whether subgroups of students exist whose empathy differs from the average, such as a group whose empathy may decline across time or a group who may remain stable, either at low or high levels of empathy. Figure 1 depicts the structural path diagram tested by LGM and GBTM. Differences in empathy between academic years were further tested by the means of paired-

sample t-tests with post hoc Bonferroni adjustment. Finally, post hoc analyses by means of regression models were carried out to determine if personality scores measured in Year 1 and medical students' endorsement of motives for studying medicine in Years 1, 4 and 5 predicted their membership in empathy trajectory groups. To avoid multi-collinearity issues, we tested separate regression models for every yearly assessment of motives, for a total of three distinct models. In every model, we used personality traits measured in Year 1 alongside the covariates age, gender and cohort membership. Motives for studying medicine measured at different assessments were added separately for every study year (i.e. Years 1, 4 and 5). Akaike's information criterion (AIC) and Bayesian information criterion (BIC) were used to compare non-nested models with lower values being indicative of a better fit (Royston, 2001). All analyses were run using Stata 15 (StataCorp. 2015. Stata Statistical Software: Release 15. College Station, TX: StataCorp LP).

Results

Preliminary analyses

We initially examined the data to detect response patterns across each self-reported scale which may indicate insufficient effort responding and result in outlier cases (Curran, 2016). Based on this initial check, all cases were retained for the analyses (Cox, 2002; Meade and Criag, 2012). The rate of missing data ranged from 0 to 11%, a percentage still considered acceptable by previous research (Peng, Harwell, Liou, and Ehman, 2006). The result of Little's test for data missing completely at random was not significant ($p = 0.776$), indicating that missing values were not dependent on any mechanism or variable. According to absolute values of skewness (values <3) and kurtosis (values <10), all psychological variables were reasonably normally distributed (see Table A1 in the Appendix) (Kline, 2015). Based on these preliminary results, the full information maximum likelihood estimation method was used in the LGM and GBTM analyses to deal with missing values. Participants

from Cohort 1 (M_{age} = 21.07, SD = 2.64) were on average significantly older than those from Cohort 2 (M_{age} = 20.38, SD = 1.59) at $p < 0.05$ (Table A2). Moreover, although effect sizes based on Cohen's *d* for comparisons of means between these groups were generally small to moderate in magnitude (effect sizes are based on Cohen's classification: 0.2 = *small*, 0.5 = *medium* and 0.8 = *large*), given further differences between the two Cohorts in terms of personality and motives, we controlled for cohort membership in the analyses where appropriate.

Latent growth modeling

Based on the preliminary analyses reported above, LGM and GBTM were run using the full information maximum likelihood estimation method to deal with missing values. LGM analyses were run for the overall sample as well as for males and for females separately, given the significant differences in empathy levels according to gender reported by previous studies (Costa et al., 2013; Hojat et al., 2002; Williams et al., 2014). LGM results reported in Table 2 (see Appendix S1 for full details on the analyses) showed that a model depicting linear growth in empathy across assessments best fitted the data for what concerned the total sample (i.e. males and females together) and females only. A no-growth model (i.e. without a slope factor) yielded the best fit for males only. The estimated mean Intercept (*M*; indicating the initial empathy level) was significant for the total sample [*M*(Intercept) = 112.86, *SE* = 0.68, $p < 0.001$)], for females [*M*(Intercept) = 115.20, *SE* = 0.72, $p < 0.001$] and for males [*M*(Intercept) = 110.93, *SE* = 1.02, $p < 0.001$]. Estimated mean slopes, representing the average degree of change in empathy across time, were not significantly different from zero for the total sample [*M*(Slope) = 0.14, *SE* = 0.15, $p = 0.380$] nor for females only [*M*(Slope) = -0.06, *SE* = 0.20, $p = 0.769$], indicating that empathy did not significantly increase or decrease over time. The best fitting model for males did not include a growth term (i.e. the slope), therefore underlying that empathy did not change also among males at the

individual level across time. The covariance between Intercept and Slope (*cov*) was not significant for the total sample [*cov*(Intercept, Slope) = 1.07, *SE* = 2.01, $p = 0.596$] nor for females only [*cov*(Intercept, Slope) = 2.24 *SE* = 2.08, $p = 0.282$], meaning that there was no significant relationship between initial empathy levels and subsequent empathy growth rates. Finally, concerning the model analysing the total sample, both the Intercept and Slope variances (*V*) were significant [*V*(Intercept) = 43.10, *SE* = 10.60, $p < 0.001$; *V*(Slope) = 1.06, *SE* = 0.63, $p < 0.001$], suggesting that there was meaningful inter-individual variability around the group average and change in empathy over time. Analyses were thus extended into GBTM to evidence whether different empathy trajectories could be observed in the current sample.

Group-based trajectory modeling

Examination of fit indexes (see Appendix S1 and Table A3) indicated that the 2-group solution was a better solution than either the 3- or 4-group ones. Figure 2 depicts JSE-S' scores across assessments for both trajectory groups. According to paired-samples *t*-tests, for students in the first group (*n* = 59; 29%; henceforth *lower and decreasing*), in Year 5 empathy was significantly lower than in Year 1 (*t* = 3.54, $p = 0.001$, Cohen's *d* = 0.49); conversely, for students in the second group (*n* = 142; 71%; henceforth *higher and stable*), empathy remained overall stable between Year 1 and Year 5 (*t* = -1.69, $p = 0.093$, Cohen's *d* = 0.15). More specifically, for *lower and decreasing* students, empathy remained stable between Year 1 and Year 4 (*t* = 1.05, $p = 0.301$, Cohen's *d* = 0.14), and between Year 4 and Year 5 (*t* = 1.56, $p = 0.126$, Cohen's *d* = 0.22), while it decreased significantly from Year 1 to Year 5 (*t* = 3.54, $p = 0.001$, Cohen's *d* = 0.49). On the other hand, for *higher and stable* students, empathy increased significantly from Year 1 to Year 4 (t = -4.85, $p < 0.001$, Cohen's *d* = 0.43) and decreased significantly from Year 4 to Year 5 (*t* = 2.66, $p = 0.009$, Cohen's *d* = 0.25) but remained stable overall between Year 1 and Year 5 (*t* = -1.69, *p* =

0.093, Cohen's $d = 0.15$). According to Bonferroni correction for multiple testing, only paired sample *t*-tests with *p*-values < 0.008 should be considered as significant, therefore the decrease in empathy for *higher and stable* students between Years 4 and 5 was classified as non-significant.

In sum, *lower and decreasing* students reported significantly lower empathy scores across all assessments ($p < 0.001$, Cohen's *d* ranged between 1.37 and 2.26) and a lower proportion of females (44%) than *higher and stable* students (62%, $p = 0.020$), while the two groups did not differ in terms of age and were composed in equal portions of students from each of the two study cohorts.

Trajectory group determinants

As an index of multi-collinearity, we considered variance inflation factors (VIF) above 4 (O'Brien, 2007). In the logistic regression models tested here, VIF ranged between 1.14 and 2.72, indicating that multi-collinearity was not an issue. Results from logistic regressions are reported in Table 2: three consecutive distinct models were tested, each one with motives measured in a different study year (Years 1, 4 and 5). Results from Model 1 showed that higher openness scores measured at Year 1 increased the probability of membership in the *higher and stable* group when including in the model motives for studying medicine measured at Year 1 (OR = 1.09, SE = 0.04, $p = 0.017$). Conversely, in Models 2 and 3, including among the predictors motives for studying medicine measured at Years 4 and 5, respectively, the effect of openness ceased to be significant while higher scores in motives pertaining to *caring for patients* (Year 4: OR = 3.78, SE = 1.33, $p < 0.001$; Year 5: OR = 2.90, SE = 0.99, $p = 0.002$) and *altruism* (Year 4: OR = 1.71, SE = 0.43, $p = 0.032$) increased the probability of belonging to the *higher and stable* group. Based on the AIC and BIC values, Model 2 seemed to provide the best fit to the data. In sum, while during preclinical years personality seems to play a slightly significant role in explaining medical students'

empathy trajectories, from the beginning of clinical years, the contribution of patients-oriented motives for studying medicine appears to become more determinant.

Discussion

The current study aimed to address both issues of between- and within-individual longitudinal changes in empathy among medical students. In accordance with Costa et al. (2013), who adopted the same methodology to assess within-individual longitudinal change in empathy, this remained stable overall from entry and throughout the initial clinical years. Neither males nor females showed significant individual rates of change across assessments. However, while this was true for the majority of the students in our current sample, confirming the results of group-based analyses carried out by previous studies (Bombeke et al., 2011; Quince et al., 2011; Roff, 2015; Thomazi et al., 2014), a smaller proportion of students exhibited a slightly decreasing trend in self-reported empathy scores. Thus, this study extended previous research findings by exploring from a person-oriented perspective the significant variability in empathy change between medical students.

Results pointing out to the personality trait openness to experience as positive predictor of empathy levels are aligned with previous studies (Costa et al., 2013; Costa et al., 2014). Such a personality dimension may indeed facilitate the establishment of good doctor/patient relationships and the ability of dealing with unexpected situations. Although agreeableness did not show any significant association with empathy trajectories, past studies have underlined the predominant role of openness over agreeableness when explaining medical students' empathy, especially adopting the NEO-FFI and JSE-S self-reported measures (Costa et al., 2013; Magalhães et al., 2012). However, the contribution of personality in explaining empathy trajectories lost significance when including motives for studying medicine during clinical years in the analyses. These results suggest that students' motives during the transition from preclinical to clinical years are particularly determinant in

explaining empathy, even in comparison with other well-known correlates of empathy, such as personality. Indeed, as evidenced by previous research, motivation is a crucial aspect of many empathy-related outcomes during clinical training, such as career choice (Lefevre et al., 2010; Okayama and Kajii, 2011) and clinical skills (Dankbaar et al., 2016). Thus, our results indicate that we may predict whether medical students will be more likely to decrease in empathy or to remain stable based on the type of motives guiding them throughout the transition into clinical training.

As Hojat et al. (2009) pointed out, for some medical students, empathy may decrease in clinical placements because of an "escalation of cynicism and atrophy of idealism" (p. 1189). Our results elaborate further on this assumption, as they showed a correspondence between specific motives for studying medicine and empathy trajectories that becomes evident especially when clinical training starts. More specifically, the first clinical year (i.e. Year 4) is where one can notice the best correspondence between altruistic and patient-oriented reasons for studying medicine with empathy trajectories. Such motives appeared to be important 'drivers' of the empathy trajectory during this initial part of training. Indeed, in this phase they may further explain why empathy increased between Years 1 and 4 among the *higher and stable* group. Conversely, *lower and decreasing* students seemed to have started their clinical training with different expectations and at this stage are partially moving away from the idealistic figure of the 'caring doctor'. The correspondence between motives for studying medicine and empathy is still partially observable during the second year of training (i.e. Year 5): here, only students with a higher inner vocation for helping patients are able to preserve empathy as it was at their entrance year. Nevertheless, even for students who have higher empathy, clinical years represent a challenging passage for their professional development. They should thus not be considered a priori as immune to stress, workload and

difficult adjustments to new situations in the clinical scenario (Dyrbye et al., 2006; Godefrooij et al., 2010; Sarikaya et al., 2006).

Educational interventions can successfully maintain and enhance empathy during medical training (Batt-Rawden et al., 2013). However, interventions should target specific groups of students on the basis of individual attributes that may alter the degree of susceptibility to different interventions so to maximise our chances of producing tangible results (Boker et al., 2004; Rosenthal et al., 2011). In this sense, our results indicate that baseline empathy scores can inform us about students' subsequent scores in the following years: those exhibiting higher initial scores are more likely to remain stable across study years whereas those exhibiting lower initial scores may be more likely to remain low or even decrease in empathy. Moreover, the associations between motives for studying medicine and empathy during clinical years should encourage us to focus our efforts during this transition and to leverage on students' inner vocation for patients' care.

In conclusion, it is important to underline the limitations of the current study. First, this research was based on a single institution; therefore, our results may not be generalizable beyond the context of the study. On a related note, our findings should be interpreted with caution in light of the small sample size we adopted. Nevertheless, the only other study – to the best of our knowledge – using LGM to examine empathy change among medical students adopted a much smaller sample (Costa et al., 2013). Indeed, as pointed out by Hamilton et al., (2003), sample size may not bias the parameter estimates to a substantive degree for this type of analysis, although samples of at least 100 are recommended (Curran et al., 2010). Secondly, the person-oriented analytical approach adopted here is exploratory by definition; therefore, more studies are needed to replicate our findings and validate the typologies of empathy trajectories we observed. In this regard, in order to better characterize empathy trajectories that may emerge among medical students, future studies should aim to cover the

entire course of undergraduate medical studies. Moreover, several other correlates of empathy, including, for example, workload and coping with stress, which are particularly relevant in the context of the passage from preclinical to clinical training (Al Kadri et al., 2011; Dahlin and Runeson, 2007), may be included in future models aimed at further describing and interpreting different trajectories of empathy among medical students. On the other hand, more effort is needed to understand the consequences of different pathways of empathy development for patient-centered healthcare outcomes and career choices among future doctors. Replications of our current results are also needed using different measures of self-reported empathy (Spatoula et al., 2019). Finally, it is crucial to point out that our findings drawing from single statements about motives for studying medicine should be taken with caution in light of the non-standardized nature of these items. Validated measures of motivation may be used to further validate the relationship between motives for studying medicine and empathy trajectories tested here and corroborate our results.

The two different empathy trajectories observed among medical students in the current sample and their correspondence with specific motives for studying medicine sustain the assumption that from the early years of medical training, students follow different pathways of professional identity development as physicians. Such a process is characterized by multiple psychological facets (Adams et al., 2006; Cohen et al., 2009; Cruess et al., 2016; Soo et al., 2016) and, for most students, involves leveraging compassionate and sensitive traits to change as a person, becoming more compassionate and sensitive so to sustain empathy (McLean, 2001). These results have important practical implications in influencing empathy development, which is a strong indicator of future professionalism and clinical competence among medical students (Boker et al., 2004; Hojat et al., 2002). Indeed, as reminded out by Dohrenwend (2018), empathy can and must be taught and it may be motivated by compassion and enhanced by intrapersonal awareness. Pedagogical strategies

focusing on students' individual differences in motives for studying medicine should thus invest in promoting internal and patients-oriented reasons for studying medicine to better train students to become professional and caring doctors in the future.

We trust future research to continue exploring the issue of longitudinal changes in empathy among medical students drawing from some of the insights of our results. Most notably, assessing empathy development from an individual- rather than group-based perspective can better describe how and if it changes during medical studies (Costa et al., 2013). More particularly, combining techniques of growth curve analysis such as LGM and GBTM we can assess whether longitudinal changes in empathy exist at the group level as well as at the individual level and integrate variable-centered and person-centered analyses (Muthén and Muthén, 2000). A person-centered focus is especially useful in empathy research within medical education where data often include heterogeneous groups of students with different aspirations and different personal values related to the medical profession. Thus, rather than focusing solely on the question of whether empathy declines or increases during medical training we can aim to understand whether this is true for all students based on the characteristics of our samples (Speelman and McGann, 2013). Finally, such an analytical approach can also be very efficient at evidencing and placing in time the determinants of specific trajectories. In this regard, our results call for a promotion of altruistic and interpersonally oriented motives, especially during clinical training when they may prove to be particularly effective at tackling the decrease in empathy among a certain portion of the medical student population.

Acknowledgments

The authors are grateful to the students who responsibly answered the questionnaires. This project has been funded by the Swiss Federal Office of Public Health and by the Faculty

of Medicine at the University of Geneva. The opinions expressed here are those of the authors and should not be attributed to these institutions.

References

Adams, K., Hean, S., Sturgis, P., & Clark, J. M. (2006). Investigating the factors influencing professional identity of first-year health and social care students. *Learning in Health and Social Care, 5*(2), 55-68.

Al Kadri, H. M., Al-Moamary, M. S., Elzubair, M., Magzoub, M. E., AlMutairi, A., Roberts, C., & van der Vleuten, C. (2011). Exploring factors affecting undergraduate medical students' study strategies in the clinical years: a qualitative study. *Advances in Health Sciences Education, 16*(5), 553-567.

Aluja, A., Garcıa, O., Rossier, J., & Garcıa, L. F. (2005). Comparison of the NEO-FFI, the NEO-FFI-R and an alternative short version of the NEO-PI-R (NEO-60) in Swiss and Spanish samples. *Personality and Individual Differences, 38*(3), 591-604.

Austin, E. J., Evans, P., Goldwater, R., & Potter, V. (2005). A preliminary study of emotional intelligence, empathy and exam performance in first year medical students. *Personality and Individual Differences, 39*(8), 1395-1405.

Batt-Rawden, S. A., Chisolm, M. S., Anton, B., & Flickinger, T. E. (2013). Teaching empathy to medical students: an updated, systematic review. *Academic Medicine, 88*(8), 1171-1177.

Boker, J. R., Shapiro, J., & Morrison, E. H. (2004). Teaching empathy to first year medical students: evaluation of an elective literature and medicine course. *Education for Health, 17*(1), 73-84.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective* (Vol. 467): John Wiley & Sons.

Bombeke, K., Van Roosbroeck, S., De Winter, B., Debaene, L., Schol, S., Van Hal, G., & Van Royen, P. (2011). Medical students trained in communication skills show a decline in patient-centred attitudes: an observational study comparing two cohorts during clinical clerkships. *Patient Education and Counseling, 84*(3), 310-318.

Brazeau, C. M., Schroeder, R., Rovi, S., & Boyd, L. (2010). Relationships between medical student burnout, empathy, and professionalism climate. *Academic Medicine, 85*(10), S33-S36.

Chen, D., Lew, R., Hershman, W., & Orlander, J. (2007). A cross-sectional measurement of medical student empathy. *Journal of General Internal Medicine, 22*(10), 1434-1438.

Chen, D. C., Kirshenbaum, D. S., Yan, J., Kirshenbaum, E., & Aseltine, R. H. (2012). Characterizing changes in student empathy throughout medical school. *Medical Teacher, 34*(4), 305-311.

Chen, D. C., Pahilan, M. E., & Orlander, J. D. (2010). Comparing a self-administered measure of empathy with observed behavior among medical students. *Journal of General Internal Medicine, 25*(3), 200-202.

Cohen, M. J., Kay, A., Youakim, J. M., & Balaicuis, J. M. (2009). Identity transformation in medical students. *The American Journal of Psychoanalysis, 69*(1), 43-52.

Corker, K. S., Donnellan, M. B., & Bowles, R. P. (2013). The development of achievement goals throughout college: Modeling stability and change. *Personality and Social Psychology Bulletin, 39*(11), 1404-1417.

Costa, M. J., Costa, P., & Magalhães, E. (2013). Response to “are personality traits really weak/moderate predictors of empathy?”. *Medical Teacher, 35*(7), 611-612.

Costa, P., Alves, R., Neto, I., Marvao, P., Portela, M., & Costa, M. J. (2014). Associations between medical student empathy and personality: a multi-institutional study. *PLoS One, 9*(3), e89254.

Costa, P., Magalhães, E., & Costa, M. J. (2013). A latent growth model suggests that empathy of medical students does not decline over time. *Advances in Health Sciences Education, 18*(3), 509-522.

Costa, P. T., & MacCrae, R. R. (1992). *Revised NEO personality inventory (NEO PI-R) and NEO five-factor inventory (NEO-FFI): Professional manual*: Psychological Assessment Resources, Incorporated.

Cox, N. J. (2002). *tsspell: Stata module for identification of spells or runs in time series*. Statistical Software Components S426901, Department of Economics, Boston College.

Crossley, M. L., & Mubarik, A. (2002). A comparative investigation of dental and medical student's motivation towards career choice. *British Dental Journal, 193*(8), 471.

Cruess, R. L., Cruess, S. R., & Steinert, Y. (2016). *Teaching Medical Professionalism: Supporting the Development of a Professional Identity*: Cambridge University Press.

Curran, P. G. (2016). Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology, 66*, 4-19.

Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognition and Development, 11*(2), 121-136.

Dahlin, M. E., & Runeson, B. (2007). Burnout and psychiatric morbidity among medical students entering clinical training: a three year prospective questionnaire and interview-based study. *BMC Medical Education, 7*(1), 6.

Dankbaar, M. E., Alsma, J., Jansen, E. E., van Merrienboer, J. J., van Saase, J. L., & Schuit, S. C. (2016). An experimental study on the effects of a simulation game on students' clinical cognitive skills and motivation. *Advances in Health Sciences Education, 21*(3), 505-521.

Dohrenwend, A. M. (2018). Defining Empathy to Better Teach, Measure, and Understand Its Impact. *Academic Medicine, 93*(12), 1754-1756.

Duffy, R. D., Manuel, R. S., Borges, N. J., & Bott, E. M. (2011). Calling, vocational development, and well being: A longitudinal study of medical students. *Journal of Vocational Behavior, 79*(2), 361-366.

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2013). *An introduction to latent variable growth curve modeling: Concepts, issues, and application*: Routledge Academic.

Dyrbye, L. N., Thomas, M. R., & Shanafelt, T. D. (2006). Systematic review of depression, anxiety, and other indicators of psychological distress among US and Canadian medical students. *Academic Medicine, 81*(4), 354-373.

Ferreira-Valente, A., Monteiro, J. S., Barbosa, R. M., Salgueira, A., Costa, P., & Costa, M. J. (2017). Clarifying changes in student empathy throughout medical school: a scoping review. *Advances in Health Sciences Education, 22*(5), 1293-1313.

Gąsiorowski, J., Rudowicz, E., & Safranow, K. (2015). Motivation towards medical career choice and future career plans of Polish medical students. *Advances in Health Sciences Education, 20*(3), 709-725.

Godefrooij, M. B., Diemers, A. D., & Scherpbier, A. J. (2010). Students' perceptions about the transition to the clinical phase of a medical curriculum with preclinical patient contacts; a focus group study. *BMC Medical Education, 10*(1), 28.

Goel, S., Angeli, F., Dhirar, N., Singla, N., & Ruwaard, D. (2018). What motivates medical students to select medical studies: a systematic literature review. *BMC Medical Education, 18*(1), 16.

Gonçalves-Pereira, M., Loureiro, J., Trancas, B., Papoila, A., & Caldas-de-Almeida, J. M. (2013). Empathy as related to motivations for medicine in a sample of first-year medical students. *Psychological Reports, 112*(1), 73-88.

Griffin, B., Bayl-Smith, P., & Hu, W. (2018). Predicting patterns of change and stability in student performance across a medical degree. *Medical education, 52*(4), 438-446.

Guilera, T., Batalla, I., Forné, C., & Soler-González, J. (2019). Empathy and big five personality model in medical students and its relationship to gender and specialty preference: a cross-sectional study. *BMC medical education, 19*(1), 57.

Hamilton, J., Gagne, P. E., & Hancock, G. R. (2003). The Effect of Sample Size on Latent Growth Models.

Hegazi, I., & Wilson, I. (2013). Maintaining empathy in medical school: it is possible. *Medical Teacher, 35*(12), 1002-1008.

Hojat, M., Gonnella, J. S., Mangione, S., Nasca, T. J., Veloski, J. J., Erdmann, J. B., . . . Magee, M. (2002). Empathy in medical students as related to academic performance, clinical competence and gender. *Medical Education, 36*(6), 522-527.

Hojat, M., Mangione, S., Nasca, T. J., Cohen, M. J., Gonnella, J. S., Erdmann, J. B., . . . Magee, M. (2001). The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educational and Psychological Measurement, 61*(2), 349-365.

Hojat, M., Mangione, S., Nasca, T. J., Rattner, S., Erdmann, J. B., Gonnella, J. S., & Magee, M. (2004). An empirical study of decline in empathy in medical school. *Medical Education, 38*(9), 934-941.

Hojat, M., Vergare, M. J., Maxwell, K., Brainard, G., Herrine, S. K., Isenberg, G. A., . . . Gonnella, J. S. (2009). The devil is in the third year: a longitudinal study of erosion of empathy in medical school. *Academic Medicine, 84*(9), 1182-1191.

Hojat, M., Zuckerman, M., Magee, M., Mangione, S., Nasca, T., Vergare, M., & Gonnella, J. S. (2005). Empathy in medical students as related to specialty interest, personality,

and perceptions of mother and father. *Personality and Individual Differences, 39*(7), 1205-1215.

Hong, M., Lee, W. H., Park, J. H., Yoon, T. Y., Moon, D. S., Lee, S. M., & Bahn, G. H. (2012). Changes of empathy in medical college and medical school students: 1-year follow up study. *BMC Medical Education, 12*(1), 122.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. London: Guilford publications.

Lefevre, J. H., Roupret, M., Kerneis, S., & Karila, L. (2010). Career choices of medical students: a national survey of 1780 students. *Medical Education, 44*(6), 603-612.

Magalhães, E., Costa, P., & Costa, M. J. (2012). Empathy of medical students and personality: evidence from the Five-Factor Model. *Medical Teacher, 34*(10), 807-812.

McLean, M. (2001). Can we relate conceptions of learning to student academic achievement? *Teaching in Higher Education, 6*(3), 399-413.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods, 17*, 437–455.

Muthén, B., & Muthén, L. K. (2000). Integrating person-centered and variable-centered analyses: Growth mixture modeling with latent trajectory classes. *Alcoholism: Clinical and experimental research, 24*(6), 882-891.

Nagin, D. (2005). *Group-based modeling of development*: Harvard University Press.

Neumann, M., Edelhäuser, F., Tauschel, D., Fischer, M. R., Wirtz, M., Woopen, C., . . . Scheffer, C. (2011). Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Academic Medicine, 86*(8), 996-1009.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*(5), 673-690.

Okayama, M., & Kajii, E. (2011). Does community-based education increase students' motivation to practice community health care?-a cross sectional study. *BMC Medical Education, 11*(1), 19.

Papageorgiou, A., Miles, S., & Fromage, M. (2018). Does medical students' empathy change during their 5-year MBBS degree? *Education for Health, 31*(3), 142-147.

Peng, C.-Y. J., Harwell, M., Liou, S.-M., & Ehman, L. H. (2006). Advances in missing data methods and implications for educational research. *Real Data Analysis, 3178*.

Piumatti, G., Abbiati, M., Baroffio, A., & Gerbase, M. W. (2018). Associations between motivational factors for studying medicine, learning approaches and empathy among medical school candidates. *Advances in Health Sciences Education*, 1-14.

Ponnamperuma, G., Yeo, S. P., & Samarasekera, D. D. (2019). Is empathy change in medical school geo-socioculturally influenced? *Medical education, 53*(7), 655–65.

Quince, T. A., Parker, R. A., Wood, D. F., & Benson, J. A. (2011). Stability of empathy among undergraduate medical students: a longitudinal study at one UK medical school. *BMC Medical Education, 11*(1), 90.

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69*(4), 617-640.

Roff, S. (2015). Reconsidering the “decline” of medical student empathy as reported in studies using the Jefferson Scale of Physician Empathy-Student version (JSPE-S). *Medical Teacher, 37*(8), 783-786.

Rosenthal, S., Howard, B., Schlussel, Y. R., Herrigel, D., Smolarz, B. G., Gable, B., ... & Kaufman, M. (2011). Humanism at heart: preserving empathy in third-year medical students. *Academic Medicine, 86*(3), 350-358.

Royston, P. (2001). Flexible parametric alternatives to the Cox model, and more. *Stata Journal, 1*(1), 1–28.

Sarikaya, O., Civaner, M., & Kalaca, S. (2006). The anxieties of medical students related to clinical training. *International Journal of Clinical Practice, 60*(11), 1414-1418.

Satterfield, J. M., & Hughes, E. (2007). Emotion skills training for medical students: a systematic review. *Medical Education, 41*(10), 935-941.

Soo, J., Brett-MacLean, P., Cave, M.-T., & Oswald, A. (2016). At the precipice: a prospective exploration of medical students' expectations of the pre-clerkship to clerkship transition. *Advances in Health Sciences Education: Theory and Practice, 21*(1), 141-162.

Spatoula, V., Panagopoulou, E., & Montgomery, A. (2019). Does empathy change during undergraduate medical education? – A meta-analysis. *Medical teacher, 41*(8), 895-904.

Speelman, C., & McGann, M. (2013). How Mean is the Mean? *Frontiers in Psychology, 4*(451).

Tavakol, S., Dennick, R., & Tavakol, M. (2011). Empathy in UK medical students: differences by gender, medical year and specialty interest. *Education for Primary Care, 22*(5), 297-303.

Thomazi, L., Moreira, F. G., & Marco, M. A. D. (2014). Assessment of fourth-year medical students' development of empathy undertaken at the federal university of são paulo (Unifesp) in 2012. *Revista Brasileira de Educação Médica, 38*(1), 87-93.

Vaglum, P., Wiers-Fenssen, F., & Ekeberg, Ø. (1999). Motivation for medical school: the relationship to gender and specialty preferences in a nationwide sample. *Medical Education*.

Ward, J., Cody, J., Schaal, M., & Hojat, M. (2012). The empathy enigma: an empirical study of decline in empathy among undergraduate nursing students. *Journal of Professional Nursing, 28*(1), 34-40.

Wellbery, C., Barjasteh, T., & Korostyshevskiy, V. (2019). Medical students' individual and social empathy: A follow-up study. *Medical teacher, 41*(6), 656-661.

Williams, B., Brown, T., Boyle, M., McKenna, L., Palermo, C., & Etherington, J. (2014). Levels of empathy in undergraduate emergency health, nursing, and midwifery students: a longitudinal study. *Advances in Medical Education and Practice, 5*, 299.

Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology, 70*(4), 767-779.

Winefield, H. R., & Chur-Hansen, A. (2000). Evaluating the outcome of communication skill teaching for entry-level medical students: does knowledge of empathy increase? *Medical Education, 34*(2), 90-94.

Zenasni, F., Boujut, E., du Vaure, B., Catu-Pinault, A., Tavani, J., Rigal, L., . . . Sultan, S. (2012). Development of a French-language version of the Jefferson Scale of Physician Empathy and association with practice characteristics and burnout in a sample of general practitioners. *International Journal of Person Centered Medicine, 2*(4), 759-766.

Table 1
Comparison of fitted latent growth models for empathy

Model	χ^2	*df*	*p*	CFI	TLI	RMSEA	$\Delta\chi^2$	Δdf	*p*(d)
Total sample (*N* = 201)									
No growth	25.24	6	< 0.001	0.86	0.93	0.13			
Linear growth	**11.82**	**3**	**0.008**	**0.94**	**0.94**	**0.12**	**13.58**	**3**	**0.004**
Nonlinear growth	10.22	2	0.006	0.94	0.91	0.14	1.60	1	0.206
Females (*n* = 114)									
No growth	26.38	6	< 0.001	0.68	0.84	0.17			
Linear growth	**6.29**	**3**	**0.098**	**0.95**	**0.95**	**0.10**	**20.09**	**3**	**< 0.001**
Nonlinear growth	5.43	2	0.066	0.95	0.92	0.12	0.86	1	0.354
Males (*n* = 87)									
No growth	**11.19**	**6**	**0.083**	**0.91**	**0.96**	**0.10**			
Linear growth	6.38	3	0.094	0.94	0.94	0.11	4.81	3	0.186
Nonlinear growth	3.82	2	0.244	0.99	0.98	0.07	2.56	1	0.110

Notes. χ^2: Chi-square; *df*: degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker–Lewis Index; RMSEA = root mean square error of approximation; $\Delta\chi^2$ = difference in likelihood ratio tests; Δdf = difference in *df*; *p*(d) = probability of the difference text. The best model solutions are marked in bold.

Table 2

Logistic regression models with personality scores and motivations for studying medicine set to predict trajectory group membership (N = 201)

Predictors	Trajectory group membership (RC: *lower and decreasing*)					
	Model 1		Model 2		Model 3	
	OR (SE) CI_{95} OR	*p*	OR (SE) CI_{95} OR	*p*	OR (SE) CI_{95} OR	*p*
Age	1.06 (0.13) 0.84, 1.36	0.587	1.05 (0.13) 0.83, 1.34	0.682	1.07 (0.11) 0.88, 1.31	0.486
Gender (RC: *Male*)	2.43 (1.15) 0.96, 6.13	0.060	1.85 (0.92) 0.70, 4.91	0.214	1.60 (0.73) 0.66, 3.89	0.300
Cohort (Ref: *Cohort 1*)	1.59 (0.67) 0.69, 3.62	0.275	2.46 (1.13) 0.99, 6.07	0.052	1.60 (0.67) 0.70, 3.64	0.261
Year 1: Neuroticism	0.98 (0.03) 0.93, 1.04	0.564	1.00 (0.03) 0.94, 1.60	0.949	0.99 (0.03) 0.94, 1.04	0.708
Year 1: Extraversion	1.06 (0.05) 0.97, 1.15	0.190	1.06 (0.05) 0.97, 1.15	0.183	1.08 (0.05) 0.99, 1.18	0.056
Year 1: Openness	**1.09 (0.04) 1.02, 1.16**	**0.017**	1.06 (0.04) 0.99, 1.15	0.076	1.07 (0.04) 0.99, 1.14	0.059
Year 1: Agreeableness	1.05 (0.04) 0.97, 1.14	0.253	1.01 (0.26) 0.93, 1.10	0.794	0.99 (0.04) 0.92, 1.08	0.885
Year 1: Conscientiousness	1.03 (0.03) 0.97, 1.10	0.371	0.99 (0.03) 0.93, 1.06	0.825	1.01 (0.03) 0.95, 1.07	0.781
Year 1: Reward	1.02 (0.19) 0.71, 1.48	0.896				
Year 1: Prestige	0.88 (0.16) 0.62, 1.25	0.487				
Year 1: Private practice	0.97 (0.14) 0.73, 1.30	0.845				
Year 1: Save lives	1.42 (0.45) 0.77, 2.64	0.263				
Year 1: Care for patients	0.68 (0.21) 0.37, 1.26	0.223				
Year 1: Altruism	1.15 (0.27) 0.72, 1.83	0.567				
Year 4: Reward			0.89 (0.21) 0.57, 1.40	0.625		
Year 4: Prestige			1.03 (0.19) 0.71, 1.49	0.150		
Year 4: Private practice			0.93 (0.16) 0.66, 1.31	0.682		
Year 4: Save lives			0.64 (0.17) 0.38, 1.07	0.089		
Year 4: Care for patients			**3.78 (1.33) 1.89, 7.53**	**<0.001**		
Year 4: Altruism			**1.71 (0.43) 1.05, 2.79**	**0.032**		
Year 5: Reward					0.75 (0.18) 0.47, 1.22	0.250
Year 5: Prestige					0.93 (0.17) 0.65, 1.33	0.690
Year 5: Private practice					0.99 (0.17) 0.71, 1.39	0.971
Year 5: Save lives					0.67 (0.16) 0.43, 1.07	0.093
Year 5: Care for patients					**2.90 (0.99) 1.49, 5.64**	**0.002**
Year 5: Altruism					1.04 (0.52) 0.65, 1.68	0.868

χ^2 (*df*), *p*	23.93 (14), 0.047	54.03 (14), <0.001	32.67 (14), 0.003
Pseudo R^2	0.125	0.256	0.157
AIC	198.153	187.165	206.053
BIC	244.559	234.637	253.439

Notes. RC: Reference category; OR: Odds Ratio; CI_{95} OR: Odds Ratio's 95% Confidence Intervals; SE: Standard Error; χ^2: Chi-square; *df*: degrees of freedom. AIC: Akaike's information criterion. BIC: Bayesian information criterion. Significant results are marked in bold.

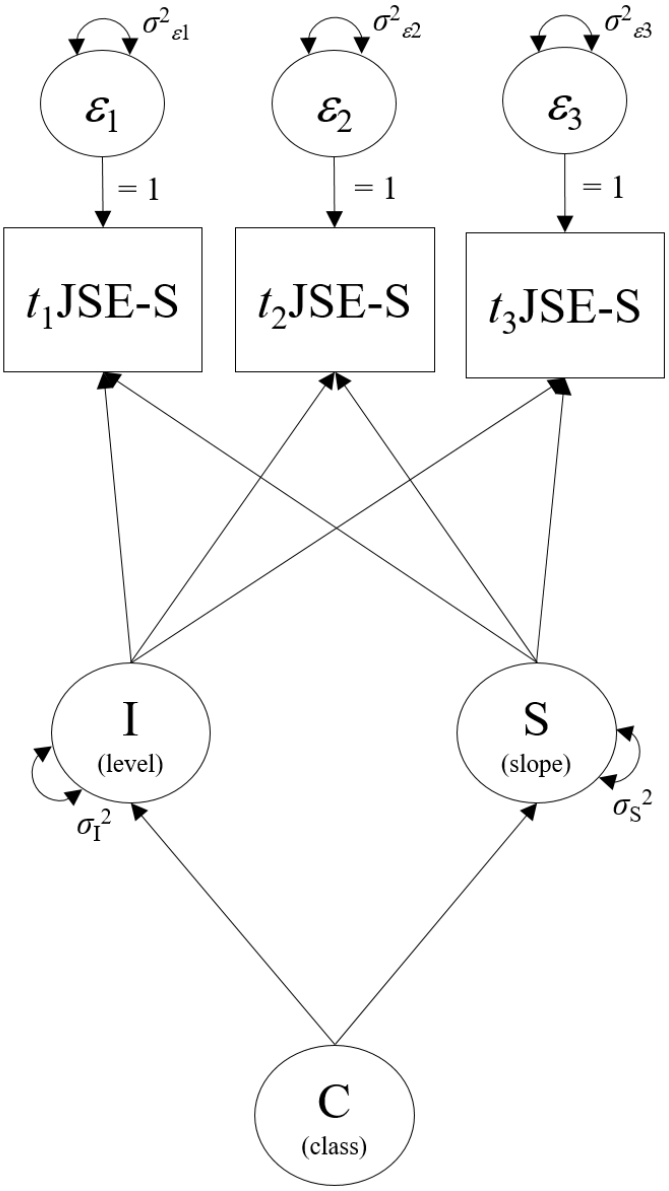

Figure 1
Group-based trajectory modeling for empathy. Observed variables are depicted as squares: t_1JSE-S, t_2JSE-S and t_3JSE-S = empathy at academic years 1, 4 and 5. Latent variables are depicted as circles: Intercept (I; average initial level), Slope (S; growth) and Class (C; latent class).

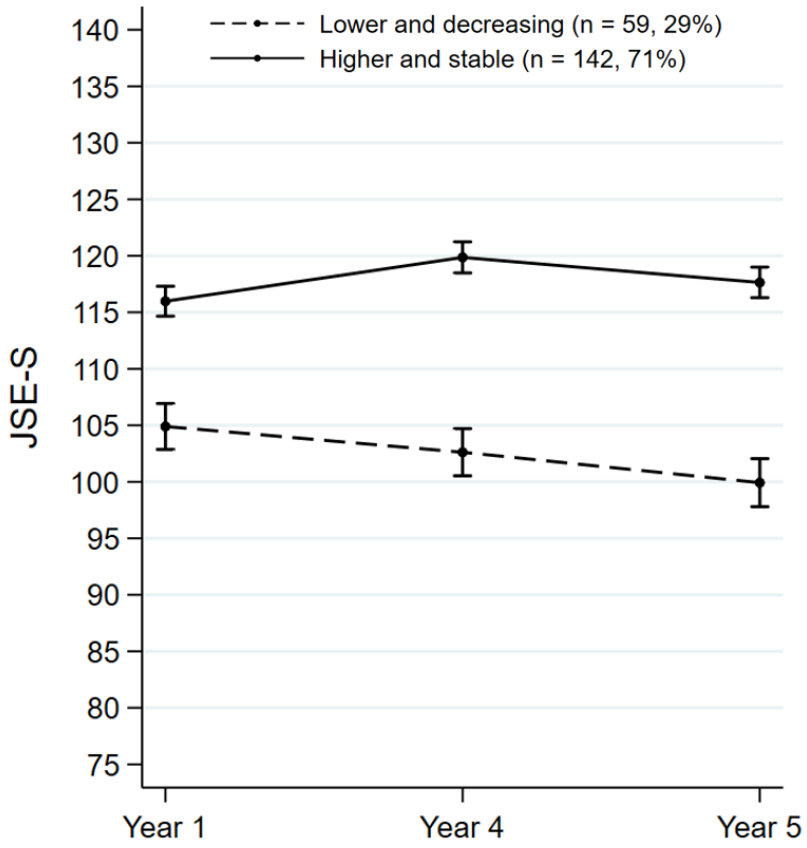

Figure 2
JSE-S' scores across assessments by trajectory groups with 95% confidence intervals
Notes. JSE-S: Student's version of the Jefferson Scale of Empathy

Empathy trajectories across medical studies: Relationships with personality and motives for studying medicine

Appendix S1

This section provides the details of the latent growth modeling (LGM) and group-based trajectory modeling (GBTM) analyses conducted using Stata 15 (StataCorp. 2015. Stata Statistical Software: Release 15. College Station, TX: StataCorp LP).

Indexes of goodness-of-fit and theoretical considerations are used to choose the best model solution for GBTM characterised by a specific number and type of trajectory groups. To determine the shape of the trajectories for empathy we followed the procedure suggested by Phan (Phan, 2011) by testing three alternative models: (1) a no-growth model where no slope component was assumed (i.e., no individual change in empathy was assumed across assessments); (2) a linear growth model assuming a linear pattern of change across assessments by fixing slope parameters to 0 at Year 1, to 4 at Year 4 and to 5 at Year 5 (following the measurement years); and (3) a nonlinear growth model where the form of change across assessments is not specified a priori and the slope parameter is fixed to 0 at Year 1 and to 5 at Year 5 to allow a separation of the intercept and slope components at baseline assessment and provide a scale of measurement for the slope. In this last model the slope parameter at Year 4 was freely estimated. Error variances for empathy at each assessment were constrained to be equal across assessments to ease model convergence (Acock, 2013; Grimm et al., 2011). Model comparison between each solution was assessed using the χ^2 difference test. The following fit indexes evaluated the overall fit of the model: the Bentler comparative fit index (CFI) (Bentler, 1990), the Tucker Lewis index (TLI) (Bentler and Bonett, 1980) and the Steiger-Lind root mean square error of approximation (RMSEA) (Steiger, 1990). Models with CFI and TLI values over 0.90 and RMSEA values

below 0.08 are considered an acceptable fit. Since the RMSEA often falsely indicated a poor fitting model when having small degree of freedoms (Kenny et al., 2015), we adopted different indexes at the same time to provide a more reliable assessment of model fit. Comparison between different LGM models are reported in Table 1 in the main text.

Once the trajectory of empathy was assessed by estimating intercept' and slope's means and variances, analyses were extended into GBTM. More specifically, the LGM analytical step tested the hypothesis of significant variance in the initial level and rate of individual change in empathy across assessments. Subsequently, GBTM tested the hypothesis that there were groups of students within the data that follow distinctive developmental trajectories (Nagin, 1999). GBTM in Stata 15 uses the maximum likelihood estimation method by a general quasi-Newton procedure (Dennis Jr et al., 1981) to predict the trajectory of each group and calculate the probability of membership in a specific group for each participants (Jones and Nagin, 2012). GBTM is an explorative approach since the number of trajectory groups are not known a priori. To determine the best solution, we looked at Bayesian information criterion (BIC) and adjusted Bayesian information criterion (aBIC) (Raftery, 1995) between alternative models. The model with the highest (least negative) value of BIC and aBIC is preferred.

As reported in Table A3, the 2-group solution characterized by a linear and a nonlinear trajectory yielded the best fit results. Both the 3-group and the 4-group models gave rise to one or more groups with a very small proportion of the observations (less than 5% of the total sample). The 2-group had an adequate proportion and sample number in each group: Group 1 – *Lower and decreasing*, 29% (n = 59); Group 2 – *Higher and stable*, 71% (*n* = 142). The average posterior probability value was 0.88 and 0.94 for each group respectively.

References

Acock, A. C. (2013). *Discovering Structural Equation Modeling Using Stata: Revised Edition*. (Stata Press)

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88*(3), 588.

Dennis Jr, J. E., Gay, D. M. & Walsh, R. E. (1981). An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software (TOMS), 7*(3), 348-368.

Grimm, K. J., Ram, N. & Hamagami, F. (2011). Nonlinear growth curves in developmental research. *Child Development, 82*(5), 1357-1371.

Jones, B. L. & Nagin, D. S. (2012). *A Stata plugin for estimating group-based trajectory models*.

Kenny, D. A., Kaniskan, B. & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research, 44*(3), 486-507.

Nagin, D. S. (1999). Analyzing developmental trajectories: a semiparametric, group-based approach. *Psychological Methods, 4*(2), 139.

Phan, H. P. (2011). Deep processing strategies and critical thinking: Developmental trajectories using latent growth analyses. *The Journal of Educational Research, 104*(4), 283-294.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111-163.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*(2), 173-180.

Table A1

Descriptive statistics for all psychological variables included in the analyses (N = 201)

Variables		Range	Mean	SD	Skewness	Kurtosis
Neuroticism	Year 1	1 – 45	21.58	8.67	-0.11	2.73
Extraversion	Year 1	15 – 44	30.14	5.45	-0.10	2.94
Openness	Year 1	14 – 43	30.25	6.35	-0.18	2.48
Agreeableness	Year 1	6 – 41	29.33	5.27	-0.67	4.44
Conscientiousness	Year 1	12 – 48	34.68	6.91	-0.47	3.46
Empathy	Year 1	77 – 140	112.68	9.55	-0.50	3.82
	Year 4	75 – 132	114.65	10.99	-0.77	3.40
	Year 5	82 – 140	112.54	11.26	-0.37	2.92
Reward	Year 1	1 – 6	3.59	1.21	-0.47	2.59
	Year 4	1 – 6	3.94	1.05	-0.88	4.14
	Year 5	1 – 6	3.80	1.11	-0.73	3.37
Prestige	Year 1	1 – 6	3.63	1.37	-0.28	2.22
	Year 4	1 – 6	3.39	1.31	-0.30	2.16
	Year 5	1 – 6	3.39	1.25	-0.20	2.30
Private practice	Year 1	1 – 6	3.78	1.39	-0.40	2.41
	Year 4	1 – 6	3.92	1.34	-0.57	2.79
	Year 5	1 – 6	4.03	1.34	-0.64	2.75
Save lives	Year 1	1 – 6	5.39	0.97	-2.55	10.95
	Year 4	1 – 6	5.10	1.12	-1.49	5.23
	Year 5	1 – 6	4.71	1.28	-0.94	3.35
Care for patients	Year 1	1 – 6	5.40	1.02	-2.41	9.48
	Year 4	1 – 6	5.42	0.83	-1.67	6.89
	Year 5	1 – 6	5.34	0.94	-1.85	7.37
Altruism	Year 1	1 – 6	4.97	1.09	-1.36	5.32
	Year 4	1 – 6	4.96	1.07	-1.05	4.05
	Year 5	1 – 6	4.84	1.09	-0.97	4.11

Table A2

Descriptive statistics for all variables included in the analyses and results of tests for significant differences between cohorts (i.e., independent samples t-tests for all continuous variables and chi-square test for gender). Values are means (standard deviations) unless stated otherwise

Variables		Cohort 1 (n = 100)	Cohort 2 (n = 101)	d	p
Females *N (%)*		57 (57)	57 (56)		0.936
Age		21.07 (2.64)	20.38 (1.59)	0.31	**0.030**
Neuroticism	Year 1	21.30 (9.02)	21.87 (8.35)	0.07	0.650
Extraversion	Year 1	32.10 (5.10)	28.18 (5.09)	0.77	**< 0.001**
Openness	Year 1	29.80 (6.26)	30.70 (6.44)	0.14	0.326
Agreeableness	Year 1	30.28 (4.62)	28.38 (5.71)	0.37	**0.012**
Conscientiousness	Year 1	35.43 (7.22)	33.93 (6.54)	0.22	0.130
Empathy	Year 1	113.06 (9.46)	112.31 (9.67)	0.08	0.582
	Year 4	114.22 (11.72)	115.10 (10.99)	0.08	0.588
	Year 5	113.30 (11.42)	111.74 (11.10)	0.14	0.350
Reward	Year 1	3.53 (1.27)	3.64 (1.15)	0.09	0.534
	Year 4	3.81 (1.11)	4.08 (0.97)	0.26	0.080
	Year 5	3.86 (1.08)	3.73 (1.14)	0.12	0.435
Prestige	Year 1	3.73 (1.37)	3.53 (1.37)	0.14	0.327
	Year 4	3.12 (1.38)	3.67 (1.18)	0.43	**0.004**
	Year 5	3.27 (1.26)	3.52 (1.23)	0.21	0.165
Private practice	Year 1	3.66 (1.50)	3.89 (1.29)	0.16	0.269
	Year 4	3.78 (1.40)	4.07 (1.27)	0.21	0.156
	Year 5	3.98 (1.45)	4.08 (1.22)	0.07	0.618
Save lives	Year 1	5.36 (1.17)	5.41 (0.75)	0.05	0.712
	Year 4	5.06 (1.15)	5.14 (1.09)	0.07	0.630
	Year 5	4.57 (1.38)	4.86 (1.16)	0.22	0.137
Care for patients	Year 1	5.40 (1.23)	5.40 (0.79)	0.01	0.976
	Year 4	5.47 (0.75)	5.37 (0.91)	0.13	0.386
	Year 5	5.30 (0.96)	5.38 (0.93)	0.08	0.568
Altruism	Year 1	5.02 (1.22)	4.93 (0.97)	0.08	0.561
	Year 4	5.06 (1.05)	4.86 (1.09)	0.20	0.188
	Year 5	4.88 (1.11)	4.80 (1.08)	0.08	0.608

Notes. Effect sizes (d) are based on Cohen's classification: 0.2 = 'small', 0.5 = 'medium' and 0.8 = 'large'. Significant p values are marked in bold.

Table A3

BIC for GBTM according to number of groups and trajectory shapes (N = 201)

Number of groups	Trajectory shapes	BIC[2]	aBIC[3]
1	0	-2125.18	-2124.16
1	1	-2128.34	-2126.80
1	2	-2125.18	-2124.16
2	0 0	-2077.18	-2075.12
2	0 1	-2079.53	-2076.96
2	0 2	-2078.03	-2074.95
2	1 1	-2079.01	-2075.93
2	**1 2**	**-2077.20**	**-2073.60**
2	2 2	-2080.26	-2076.16
3	0 0 0*	-2067.50	-2064.42
4	0 0 0 0*	-2072.37	-2068.27

Notes. GBTM: Group based trajectory modeling. Trajectory shapes: 0 = zero-order; 1 = linear; 2 = nonlinear. The chosen solution is marked in bold.

[2]BIC = Bayesian information criterion (for the total number of observations)

[3]aBIC = Bayesian information criterion (for the total number of participants)

*One or more of the groups had a very small proportion of the observations.