Journal of Educational Evaluation for Health Professions

J Educ Eval Health Prof 2014, 11: 17 • http://dx.doi.org/10.3352/jeehp.2014.11.17

Open Access

eISSN: 1975-5937

REVIEW ARTICLE

Confounding factors in using upward feedback to assess the quality of medical training: a systematic review

Anli Yue Zhou[1]*, Paul Baker[2]

[1]Royal Bolton Hospital, Lancashire, United Kingdom; [2]North Western Deanery, Manchester, United Kingdom

Abstract

Purpose: Upward feedback is becoming more widely used in medical training as a means of quality control. Multiple biases exist, thus the accuracy of upward feedback is debatable. This study aims to identify factors that could influence upward feedback, especially in medical training. **Methods:** A systematic review using a structured search strategy was performed. Thirty-five databases were searched. Results were reviewed and relevant abstracts were shortlisted. All studies in English, both medical and non-medical literature, were included. A simple pro-forma was used initially to identify the pertinent areas of upward feedback, so that a focused pro-forma could be designed for data extraction. **Results:** A total of 204 articles were reviewed. Most studies on upward feedback bias were evaluative studies and only covered Kirkpatrick level 1-reaction. Most studies evaluated trainers or training, were used for formative purposes and presented quantitative data. Accountability and confidentiality were the most common overt biases, whereas method of feedback was the most commonly implied bias within articles. **Conclusion:** Although different types of bias do exist, upward feedback does have a role in evaluating medical training. Accountability and confidentiality were the most common biases. Further research is required to evaluate which types of bias are associated with specific survey characteristics and which are potentially modifiable.

Key Words: *Bias; Confidentiality; Feedback; Quality control; Social responsibility*

INTRODUCTION

Multiple methods of feedback exist, which include downward feedback, upward feedback, peer feedback and self-evaluation. The most commonly known form of feedback is downward appraisal, where the supervisor gives feedback to the subordinate [1]. However, upward feedback, where the feedback is given from the subordinate to the supervisor is becoming more recognized and adopted, especially in the private sector. It has been reported that over 90% of fortune 100 companies in the United States participate in some form of upward feedback [1]. The role of upward feedback has also been widely acknowledged within the educational sector as well, where students give feedback to their lecturers [2-7]. Within medical training, the General Medical Council (GMC) in the United Kingdom has adopted upward feedback to monitor teaching performance for quality control purposes [8]. Although upward feedback has been advocated by the GMC, it is not immune from bias and there has been much debate about the accuracy of upward feedback [9-17]. This systemic review has been prompted by the increasing significant role of upward feedback as medical training becomes more closely regulated. Bias present within upward feedback could potentially skew feedback on medical training and this review aims to identify these factors.

*Corresponding email: Anlizhou89@doctors.org.uk

Received: April 16, 2014; Accepted: August 12, 2014;
Published: August 13, 2014
This article is available from: http://jeehp.org/

METHODS

Search strategy

In order to obtain a comprehensive overview of the literature in upward feedback, a total of 35 databases were searched (Embase, Medline, PsychINFO, Cochrane and EBM Reviews, Allied and Complementary Medicine, CAB and ATLA Religion Database, Econ lit, GeoBase, Global Health, Health and Psychosocial Instruments, HMIC Health and Management, Index to Foreign Legal Periodicals, International Pharmaceutical Abstracts, Maternity and Infant Care, The Philosopher's Index, Social Policy And Practice, Zoological Records, BNI, CINAHL, Health Business Elite, ERIC, British Educational Index, ASSIA, Web of Knowledge, Social Care Online, Sage Full Text Journals, IBBS, National Research Register Archive, Proquest, Wiley Online Library, Taylor and Francis, Engineering Village, Scopus, Science Direct, PubMed). A stratified search involving multiple keywords was used (Fig. 1).

Searches were initially done to search within all fields. If more than 1,000 results were returned, then the search would be repeated to search within keywords, then abstract and then within the title in order to narrow down results to less than 1,000 articles. Search results of less than 1,000 articles were reviewed by reading the abstract; relevant abstracts were then shortlisted. If no abstract was available, but the title appeared relevant, this would also be temporarily shortlisted until further information could be obtained from the full article. Further references were found by reviewing the reference bibliography of the shortlisted articles.

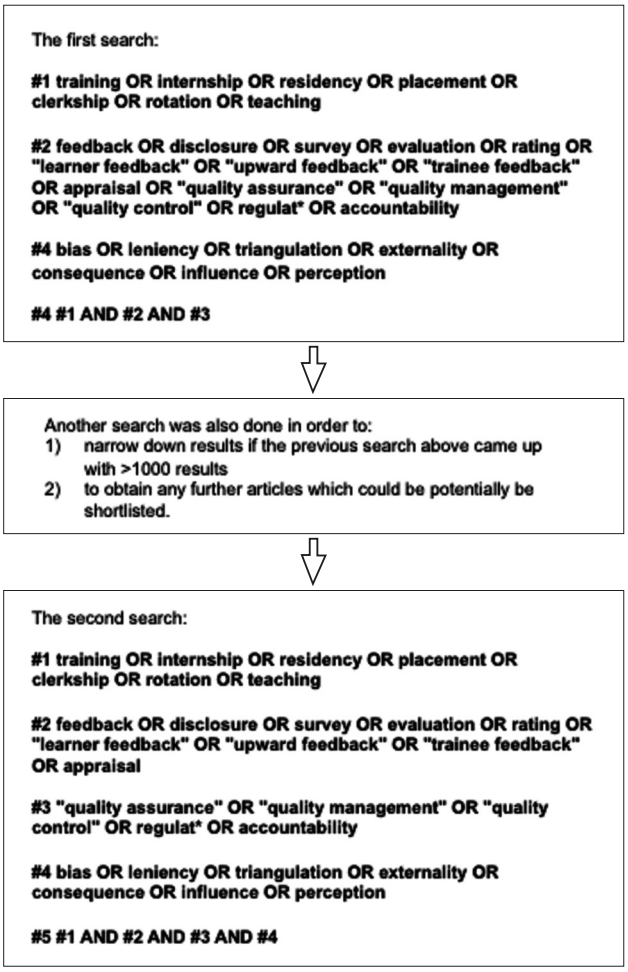

Fig. 1. Search strategy of related papers for systemic review.

Inclusion and exclusion criteria

Both medical and non-medical articles written in English were included. No time limit was set. Books were excluded from the search.

Data management techniques

A proforma was developed to allow efficient and relevant data extraction. This included: study method (e.g., observational or review article), profession, type of participant, geographical location, purpose of feedback (e.g., summative or formative), feedback subject (e.g., trainer, training or environment), qualitative/quantitative feedback, the use of controls and type of intervention involved (e.g., counseling, timing of feedback), type of feedback used (e.g., paper survey, semi-structured interviews), quality of questions (e.g., closed, open), duration of study, number of participants, response rates, types of bias present (overt and implied), Kirkpatrick level [18] and whether outcomes were addressed.

RESULTS

Literature search and selection

A total of 8,914 potential articles were found using the search strategy (Fig. 1), in which 291 articles were shortlisted. The shortlisted articles were then subsequently pooled together and duplicates were removed. A total of 169 articles were shortlisted after the removal of duplicates. By reviewing the reference bibliography of the shortlisted articles, a further 70 articles were shortlisted. A total of 239 references were shortlisted. After reviewing the articles 35 articles were excluded from further analysis. This was due to: 10 articles were not relevant to the objective, 1 reference was a book, complete versions were not obtainable for 21 references, 2 references were not written in English and 1 reference was a duplicate of another shortlisted reference but was under a different title. This lead to a total of 204 articles being analyzed, all of which are presented in Table 1.

Table 1. Summary of all the references shortlisted and analysed in this systematic review

Type of participant	Medical	Non-medical
Undergraduate	Langenfeld et al. [50], Rabow et al. [84], Brasher et al. [100], Metcalfe and Matharu [106], Blue et al. [117], Iqbal and Khizar [119], Solomon et al. [127], Johnson and Chen[129], Windish et al. [130], Ramsey et al. [135], Tochel et al. [144], Fallon et al. [152], Stritter et al. [153], Shellenberger and Mahan [154], Cohen et al. [155], Dolmans et al. [156], Donnelly and Wooliscroft [157], Irby and Rakeshaw [158], Parikh et al. [159], Wilson [160], De et al. [161], Duffield and Spencer [162], Tiberius et al. [163], Gil et al. [164], Pfeifer and Peterson [165]	Al issa and Sulieman [9], Bernardin [19], Crittenden and Norr [20], Adams and Umbach [21], Wolbring [22], Remedios and Lieberman [24], Chen and Hoshower [25], Worthington [26], Kember and Wong [27], Marsh [28], Marsh [29], Marsh and Roche [30], Rowden and Carlson [31], Goos et al. [32], Davies et al. [33], Blackhart et al. [34], Dwinell and Higbee [35], Burdsal and Bardo [36], Theall and Franklin [38], Feldman [39], Sojka et al. [40], Berk [41], Greenwald and Gillmore [42], Gigliotti and Buchtel [43], Doyle and Crichton [44], Aleamoni [46], Kember and Leung [52], Roch and McNall [67], Atwater et al. [73], Redman and McElwee [74], Chan and Ip [76], Henderson et al. [77], Brugnolli et al. [78], Midgley [80], Per Palmgren [82], Olson et al. [87], Braine and Parnell [88], Perli and Brugnolli [89], Heffernan et al. [90], Kelly [91], El Ansari and Oskrochi [102], Berber [110], Robbins and DeNisi [134], Govaerts et al. [146], Surratt and Desselle [149], Cardy and Dobbins [166], Henzi et al. [167], Parker and Carlisle [168], Cooke et al. [169], Myall et al. [170]
Postgraduate	Archer et al. [10], Barrow and Baker [12], Coats and Burd [13], Arah et al. [47], Schneider et al. [51], Scott et al. [53], Ahearn et al. [55], Grava-Gubins and Scott [63], Owen [64], Fiander [65], Risucci et al. [66], Kolarik et al. [85], Smith et al. [86], Ranse and Grealish [92], O'Connor et al. [95], Luks et al. [96], Turnball et al. [97], Biller et al. [98], Carpenter et al. [99], Busari et al. [101], Basu et al. [103], Whang et al. [104], Devlin et al. [105], Barrett et al. [107], Steiner et al. [108], Getz and Evens [109], Girard et al. [111], Antiel et al. [112], Lin et al. [113], Ratanawongsa et al. [114], Thangaratinam et al. [115], Kanashiro et al. [116], Watling et al. [118], Watling et al. [120], Pearce et al. [122], Conigliaro et al. [123], Dech et al. [124], Yarris et al. [125], Sargeant et al. [126], Sender Lieberman et al. [128], Tortolani et al. [131], O'Brien et al. [132], Claridge et al. [133], Hayward et al. [136], Sargeant et al. [137], Paice et al. [141], Ryland et al. [142], Rose et al. [145], Bing-you et al. [151], Kjaer et al. [171], Hrisos et al. [172], Beckman et al. [173], Mattern et al. [174], Kendrick et al. [175], Keitz et al. [176], Moalem et al. [177], Sargeant et al. [178], Schuh et al. [179], Vasudev et al. [180], Ellrodt [181], Harrison and Allen [182], Dola et al. [183], Cohn et al. [184], Fisher et al. [185], Pankhania et al. [186], Welch et al. [187], Greysen et al. [188], Mailloux [189], Buschbacher and Braddom [190], Cooke and Hutchinson [191], Holland et al. [192], Sabey and Harris [193], Nettleton et al. [194], Chamberlain and Nisker [195], Verhulst and Distlehorst [196], Guyatt et al. [197], Barclay et al. [198]	McCarthy and Garavan [1], Hall et al. [14], Caskie et al. [15], Smith and Fortunato [16], Kudisch et al. [17], Mullen and Tallant-Runnels [37], Tews and Tracey [48], Tews and Tracey [49], Smither et al. [54], Antonioni and Park [56], Tsui and Barry [57], Ryan et al. [59], Antonioni [61], Goodwin and Yeo [62], Antonioni [68], Bettenhausen and Fedor [69], Westerman and Rosse [70], Mathews and Redman [71], Reid and Levy [72], Redman and Snape [75], Cohan [81], Raikkonen et al. [83], Beecroft et al. [93], Sit et al. [94], Brett and Atwater [138], Barclay et al. [139], Tourish and Robson [148], Dipboye and de Pontbriand [199], Copp et al. [200], Bratt and Feizer [201], Smither and Walker [202], Becker et al. [203]
Both undergraduate and postgraduate	Gross et al. [23], Schum et al. [45], Albanese [58], Eva et al. [60], Cannon et al. [121], Irby [140], Watling and Lingard [143], Williams et al. [147], Mcleod et al. [204], Bennett et al. [205]	Ilgen et al. [150], Henzi et al. [206], Henzi et al. [207]

Demographics

More than 50% of the references were related to the medical profession (n = 109). Other professions that have commonly utilized upward feedback include teaching and education (n = 39), nursing (n = 22) and management (n = 18). The majority of references included postgraduate participants (n = 106). Thirteen references included both undergraduate and postgraduate participants. A large proportion of references were from North America (Fig. 2).

Types of studies and feedback

Studies were categorized according to the definitions in Table 2. Most references were evaluation studies (n = 176) and most studies were done for formative purposes (n = 172). A large majority of studies were quantitative (n = 152) and high proportion of studies used paper surveys as a means of evaluating upward feedback (n = 124). Most studies (n = 162) only covered Kirkpatrick level 1, reaction. The median response rate was 76%, the median number of participants was 198 and the median duration of the study was 6 months. Only 1/3 of references addressed the outcomes of their study by develop-

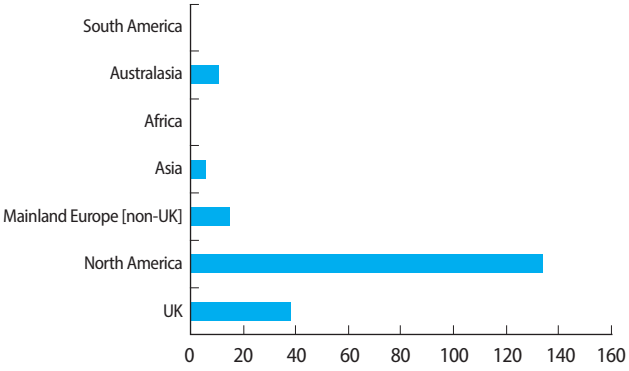

Fig. 2. Geographical locations of studies in the targeted papers for systemic review.

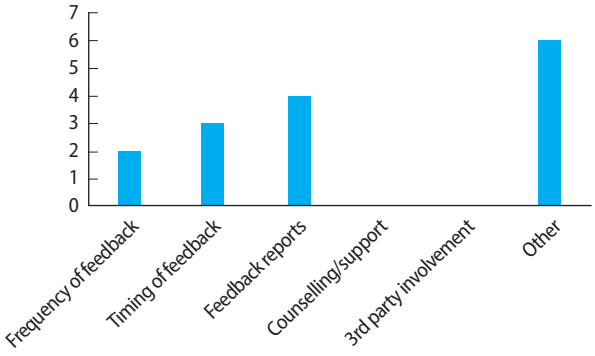

Fig. 3. Type of interventions used in control studies.

Table 2. Summary of categories used within the focused proforma

Proforma categories	Further information
1. Number	Each article was allocated a number to allow easy identification.
2. Study method	What type of study was it?
3. Profession	What profession were the participants?
4. Type of participant	Undergraduate or postgraduate or both?
5. Geographical location	Which continent was the article from?
6. Purpose of study	Was the study for summative (for promotional/reward purposes) or formative (for improvement/development) purposes?
7. Feedback subject	Feedback on training, trainer or learning environment?
8. Quality of feedback	Quantitative or qualitative?
9a. Were controls used?	Controls may be used to compare the efficacy of different interventions.
9b. Type of interventions	
10a. Type of evaluation	What type of feedback method was used? e.g., paper survey, focus groups
10b. Quality of questions	What types of questions were used? e.g., closed, open mixture
11. Duration of study	Measured in months
12. Number of participants	Total number of participants giving upward feedback
13. Response rates	Measured in percentages
14. Types of bias	Split into implied and overt: Overt bias would be explicitly mentioned by the authors within the study. Implied bias would be identified by the reviewer as potential bias but was not mentioned within the study.
15. Action plans	Did the authors address the outcomes/consequences of the article? Was an action plan devised to address this?
16. Kirkpatrick levels	Which level? [18] (1) Reaction: What do the raters think about their trainer/training/environment? (2) Learning: Was the ratee able to learn from this feedback? This can be identified through mechanisms such as feedback reports, receiving results. (3) Behavior: Did the ratee change their behavior due to this feedback? This can be reflected in repeat ratings. (4) Results: Was there any improvement in teaching after receiving the feedback? Did others benefit from this improvement? For example, did exam rates improve? Did this change improve company profits?

ing an action plan. Furthermore, only 11 studies used controls to compare different interventions (Fig. 3).

Types of bias

Types of bias data separated into implied and overt bias. Implied bias involves factors that potentially could affect the upward feedback process but was not explicitly acknowledged within the article. Overt bias included factors affecting the upward feedback process that were mentioned within the article. A summary of the different types of bias found in this systematic review can be found in Table 3. Accountability and confidentiality were the most common biases recognized within references. On the other hand, the method of feedback, which involves the type of survey, the location, the use and methodology of reminders and the duration, were most commonly implied within articles but not explicitly acknowledged (Table 4).

Table 3. Different types of bias identified within the systematic review

Type of bias	Further information
1. Affect/leader-member relationship	Defines the relationship between ratee and rater [57,134]. The bias of liking someone may lead to potentially inaccurate ratings.
2. Motivation	Low response rates may not be representative of the sampled population. This could potentially be due to lack of motivation. Prior interests, including prior subject interest [4,30] could also affect participation and enthusiasm. For example, did students volunteer themselves to enter into the study? A response rate of 60% or more is perceived as an acceptable level [208]. Articles that explicitly mention rater motivations, enthusiasm or prior subject interests were also included.
3. Fear and retaliation, career progression	The fear that honest ratings could lead to retaliation and affect career progression, could potentially affect upward feedback outcomes [12].
4. Self efficacy, lack of understanding/knowledge of upward feedback, role appropriateness	Do raters feel they are suitable/appropriate/confidence to rate their superiors [11,17]?
5. Cynicism and trust, perceived usefulness	Raters may not feel their voice will be heard and may be skeptical that changes will be made according to their feedback [16].
6. Ingratiation, yea saying, leniency, reward anticipation/incentives	Raters may rate leniently as a means of showing ingratiation or to receive reward in return [11].
7. Method of feedback	This includes how survey was implemented e.g paper, online, the location of survey implementation [115], whether any reminders and method of reminders [55]. Also included whether the survey was done over a period of time or only used 1 day/session [115].
8. Voluntary/compulsory	All members had to participate or could choose not to participate.
9. Frequency/timing, opportunity to observe	The timing of the survey: Was it done straight after rotation, or done many months after rotation, or done in the middle of the rotation [201].
10. Cultural/gender	Cultural differences may affect survey accuracy [78,119]. Gender could affect survey differences e.g., nursing where the survey population is predominantly female [83].
11. Halo effect	Raters have a tendency to give similar ratings to all aspects of a survey [11,57]. Raters are not able to differentiate between different traits.
12. End aversion/extreme response	End aversion: the avoidance of extreme ratings [11]. Extreme response: always rating very high/very low scores [11].
13. Survey fatigue	If there are multiple surveys to complete in the study or if the survey was very long, then this could affect survey accuracy.
14. Survey purpose	Was the survey for administrative or developmental purposes [11,41]? Why was the survey done?
15. Others	Potential biases that could also potentially affect bias but not mentioned above. e.g., recall bias [201].

Table 4. Summary of types of upward feedback bias identified

Type of feedback bias	Implied	Overt
Affect, leader-member relationship	76	39
Motivation	42	14
Fear and retaliation	31	32
Self efficacy, lack of understanding/knowledge of upward feedback, role appropriateness	56	28
Cynicism and trust, perceived usefulness	67	32
Accountability and confidentiality	54	117
Ingratiation, yeah saying, leniency, reward anticipation/incentives	30	52
Method of feedback	104	39
Voluntary/compulsory	35	102
Frequency/timing opportunity to observe	37	31
Cultural or gender bias	68	23
Halo effect	8	10
End aversion/extreme response	14	5
Survey fatigue	50	8
Survey purpose	66	37
Others	13	11

DISCUSSION

This review shows that multiple sources of bias, in the important task of using feedback in the assessment of training quality, are already described.

Feedback philosophy

Although there has been extensive research on upward feedback within an undergraduate classroom setting [2-7,9,19-46], the high proportion of references related to the medical profession and to postgraduate participants confirms the popularity of upward feedback in postgraduate medical training. The majority used surveys for formative purposes, which can provide the trainer/teacher with guidance on their current performance. The lack of studies for summative purposes could be due to raters tending to be over-lenient when upward feedback was for administrative purposes [14,17,39]. However, in contrast, Smith and Fortunato [16] found that rating purpose did not affect intentions to provide honest ratings since raters

may use the purpose as a tool to retaliate and reward their supervisors. Upward feedback could potentially be used as a tool to develop clinical trainers and to give guidance to clinical educators on their own career plans [47]. However, the effectiveness of upward feedback could be confounded by multiple factors, which will be discussed below. Most studies only evaluated Kirkpatrick level 1–reaction, which mostly involved surveying subordinate's views on certain topics. Only 10 studies covered Kirkpatrick level 4–outcomes [1,4,5,38,44,48-52]. The majority of studies did not address the consequences or results of the study. This could be because it may be difficult to develop specific action plans based on Kirkpatrick level-one evidence. Furthermore, very few studies specifically compare the different factors or their effect on feedback quality.

Study administration

Upward feedback usually involves subordinates to appraising their superiors or training, hence it is not surprising that the majority of studies were evaluation studies. Only one study was a randomized controlled trial that stratified participants into 3 groups (online survey, simultaneous paper and online survey, sequential online and paper survey) [53]. This study found that the sequential survey method, in which online and paper surveys were administered at different times, gave the highest response rate but increased costs [53]. The small number of studies involving controls could be due to time and financial constraints. Controlled trials of educational interventions are rare, but more studies may need to include controls if we are to assess the efficacy of the different interventions. Without evidence for the effectiveness of interventions, it may be difficult for trainers to accept upward feedback from their subordinates. Tews and Tracey [49] showed that managers who participated either in self-coaching courses or upward feedback intervention, improved interpersonal scores compared to controls. Managers who participated in the upward feedback training scored higher overall [49]. This could due to the fact that upward feedback, if utilized appropriately, can facilitate information sharing, act as a refresher in order to avoid complacency and promote further development of skills [48]. Another form of support in upward feedback was the use of feedback reports, as demonstrated in Smither et al. [54]'s study. Feedback reports enabled managers to improve their managerial skills and also encouraged communication with their subordinates. However, adequate support with regular formal feedback in order to facilitate the process [48], may be difficult to orchestrate in medical training where clinical educators work shift patterns. Moreover, the costs of facilitating upward feedback support may be quite high.

It is only in recent years as the internet has become widely accessible that online surveys have become more commonly utilized, hence why paper surveys were still the most commonly used form of feedback method within this review. Online surveys are cheaper and easier to administer in comparison to paper surveys and allow people to do the survey at a time that is convenient for them [55]. Scott et al. [53]'s study showed that although doctors in training did not give the highest response rates overall, trainee doctors gave the highest response rate when the survey was online. This may suggest the increasing role of online surveys in the newer generation of doctors. Furthermore, using online surveys to monitor training and trainers could allow the data to be more representative of the population of doctors in training.

Human factors in upward feedback bias

Affect describes the feeling of liking someone [56,57]. It has been thought that affect can lead to leniency because it can prevent one's ability to objectively and rationally evaluate someone [58]. Al-issa found that students gave higher ratings to teachers who they got along with [9]. Moreover, Antonioni and Park showed that the leniency was more profound in both peer and upward feedback compared to downward feedback [56], suggesting that affect may play a role in both peer and upward feedback. In contrast, a study by Ryan et al. [59] found that recipients of feedback were more likely to accept feedback from those who they are already acquainted to and this finding was confirmed in another study [60]. This could suggest that supervisors may be more accepting of honest feedback and this may encourage subordinates who have a positive relationship with their supervisors to give honest feedback.

Antonioni [61] found that participants who were not anonymous when they gave upward feedback did give higher ratings compared to anonymous participants. Furthermore, fewer participants stayed in the study after finding out they were in the group which could be identified [61]. However, this study was implemented within an insurance company where upward feedback could potentially be for used for summative purposes. This could lead to greater inflation in order to minimize the negative consequences. In contrast, upward feedback in medical training is more likely to be for formative purposes in order to further develop the clinical educator. Many studies have allowed upward feedback response to be confidential due to potential rating inflation [3,4,7,12-15,17,22-24,26,28,34-39,43-45,47-50,52-55,57,58,61-142], hence accountability and confidentiality was the most commonly acknowledged type of bias found within this systematic review. In contrast, Roch and McNall [67] that investigated whether anonymity affected ratings found that students who were not anonymous actually gave lower ratings compared to anonymous raters. Non-anonymous raters may feel more pressure to give high quality ratings [67]. So, there still may be a role for surveys in which sub-

ordinates may be accountable for their ratings. Furthermore, supervisors seem to be more accepting of accountable surveys [61]. Unfortunately, in potentially negative situations, anonymity seems likely to be the best policy.

Reward anticipation could be related to evaluation inflation. Previous studies have found that course grades can significantly predict student ratings [7,9], but the causation is unclear. Marsh and Roche [30] found that giving high grades were not related to higher student evaluation, but instead a lot of the variation within student evaluations could be accounted for by prior subject interest, higher and challenging workloads and learning. Furthermore, Abrami et al. [6] found that student grades were unlikely to have an effect on student ratings. The relationship of reward and ratings has been inconsistent and can be subjected to interpretation, hence the need for further research in this area.

Even if confidentiality concerns are addressed, this may still affect participation due to fear and retaliation [10,12,15,61,62,132]. The miscorrelation of self-perception and upward feedback results could affect acceptability and credibility of upward feedback since it threatens self-esteem [143]. Multiple factors can affect people's receptivity to feedback, this includes their motivation, fear and expectations [60]. However, if feedback is delivered appropriately and is perceived as valuable, then this can minimise the risk of negative emotions and dismissal of the feedback [60]. This is likely to require specialist input e.g., counseling which may have extra cost implications.

A lack of trust and cynicism was not an uncommon finding in both medical [45,53,55,58,137,142,144-147] and non-medical feedback [5,9,15-17,21,26,38-40,52,61,67,70,71,75 81,82,91,148-150]. If there is discrepancy between self-ratings and upward feedback ratings [128,145], there is a possibility that the recipient may not find the feedback credible. Also poorly designed surveys that may lack useful feedback can lead to reluctance to change. Even trainees question the credibility of some of the feedback provided by their supervisors [151], hence it is likely that supervisors may do the same to feedback from trainees. Moreover, upward feedback, especially in an undergraduate setting has been compared to 'popularity contests.' Aleamoni [46]'s review article demonstrated that evidence supports the fact that students are able to judge the effectiveness of teaching. However, attitudes are harder to modify and this misperception may still lead to faculty being more resistant to change. This resistance could in turn affect raters' enthusiasm, especially if previous experiences of upward feedback lead to no improvement.

Limitations

Although a comprehensive search was done, however, this may not be representative of all the data available on upward feedback. Also, a total of 35 articles shortlisted in the systematic review were not included in the results. There could potentially be other types of bias present in literature that was not reviewed within this systematic review. Moreover, we have identified a number of different biases that are involved in upward feedback, however we have not investigated how these biases can be minimised. Further research will be required in order to determine whether these biases are interrelated and if it is possible to minimise the effects of different biases, especially human factors.

CONCLUSION

Upward feedback is a multidimensional form of feedback that can lead to improvement if facilitated and implemented appropriately. This systematic review has shown that multiple different types of bias can exist within upward feedback. The established literature acknowledges and suggests likely causes of bias, without thoroughly investigating their effect on feedback quality. This highlights the importance for managers of training to consider important factors such as survey method and intended uses when designing and interpreting feedback. Currently, a mixed approach with triangulation of methods seems to be the best way to evaluate medical training. Further research is required in order to evaluate which types of bias are associated with specific survey characteristics and which factors are potentially modifiable.

ORCID: Anli Yue Zhou: http://orcid.org/0000-0003-2402-8289; Paul Baker: http://orcid.org/0000-0003-1493-9660.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

SUPPLEMENTARY MATERIAL

Audio recording of the abstract.

REFERENCES

1. McCarthy AM, Garavan TN. 360 degrees feedback process: performance, improvement and employee career development. J Eur Ind Train. 2001;25:5-32. http://dx.doi.org/10.1108/03090590110380614
2. Marsh HW, Roche L. The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness. Am Educ Res J. 1993;30:217-251. http://dx.doi.org/10.3102/00028312030001217

3. Kogan LR, Schoenfeld-Tacher R, Hellyer PW. Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. Teach High Educs. 2010;15:623-636. http://dx.doi.org/10.1080/13562517.2010.491911
4. Marsh HW, Roche L. Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias, and utility. Am Psychol. 1997;52:1187-1197. http://dx.doi.org/10.1037/0003-066X.52.11.1187
5. Marsh HW. Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and utility. J Educ Psychol. 1984;76:707-754. http://dx.doi.org/10.1037/0022-0663.76.5.707
6. Abrami PC, Dickens WJ, Perry RP, Leventhal L. Do teaching standards for assigning grades affect student evaluations of teaching? J Educ Psychol. 1980;72:107-118. http://dx.doi.org/10.1037/0022-0663.72.1.107
7. Brockx B, Spooren P, Mortelmans D. Taking the grading leniency story to the edge: the influence of student, teacher, and course characteristics on student evaluations of teaching in higher education. Educ Assess Eval Account. 2011;23:289-306. http://dx.doi.org/10.1007/s11092-011-9126-2
8. General Medical Council. The GMC quality framework for speciality including GP training in the UK. General Medical Council; 2010 [cited 2014 Apr 23]. Available from: http://www.gmc-uk.org/6___PMETB_Merger___Governance_Standards_and_Policies___Annex_D.pdf_36036849.pdf
9. Al-Issa A, Sulieman H. Student evaluations of teaching: perceptions and biasing factors. Qual Assur Educ. 2007;15:302-317. http://dx.doi.org/10.1108/09684880710773183
10. Archer J, McGraw M, Davies H. Assuring validity of multisource feedback in a national programme. Arch Dis Child. 2010;95:330-335. http://dx.doi.org/10.1136/adc.2008.146209
11. Berk RA, Naumann PL, Appling SE. Beyond student ratings: peer observation of classroom and clinical teaching. Int J Nurs Educ Scholarsh. 2004;1:1-26. http://dx.doi.org/10.2202/1548-923x.1024
12. Barrow P, Baker P. Factors that affect upward feedback in general surgery registrar training. 2013 (unpublished data).
13. Coats RD, Burd RS. Intraoperative communication of residents with faculty: perception versus reality. J Surg Res. 2002;104:40-45. http://dx.doi.org/10.1006/jsre.2002.6402
14. Hall JL, Leidecker JK, DiMarco C. What we know about upward appraisals of management: facilitating the future use of UPAs. Hum Resour Dev Q. 1996;7:209-226. http://dx.doi.org/10.1002/hrdq.3920070303
15. Mehr K, Ladany N, Caskie G. Trainee nondisclosure in supervision: what are they not telling you? Couns Psychother Res. 2010;10:103-113. http://dx.doi.org/10.1080/14733141003712301
16. Smith AF, Fortunato VJ. Factors influencing employee intentions to provide honest upward feedback ratings. J Bus Psychol. 2008;22:191-207. http://dx.doi.org/10.1007/s10869-008-9070-4
17. Kudisch JD, Fortunato VJ, Smith AF. Contextual and individual difference factors predicting individuals' desire to provide upward feedback. Group Organ Manage. 2006;31:503-529. http://dx.doi.org/10.1177/1059601106286888
18. Kirkpatrick DL. Techniques for evaluating training programs. Train Dev J. 1979;33:78-92.
19. Bernardin JH. Effects of rater training on leniency and halo errors in student ratings of instructors. J Appl Psychol. 1978;63:301-308. http://dx.doi.org/10.1037/0021-9010.63.3.301
20. Crittenden KS, Norr JL. Student values and teacher evaluation: a problem in person perception. Sociometry. 1973;36:143-151. http://dx.doi.org/10.2307/2786563
21. Adams MJ, Umbach PD. Nonresponse and online student evaluations of teaching: understanding the influence of salience, fatigue, and academic environments. Res High Educ. 2011;53:576-591. http://dx.doi.org/10.1007/s11162-011-9240-5
22. Wolbring T. Class attendance and students' evaluations of teaching: do no-shows bias course ratings and rankings? Eval Rev. 2012;36:72-96. http://dx.doi.org/10.1177/0193841X12441355
23. Gross J, Lakey B, Edinger K, Orehek E, Heffron D. Person perception in the college classroom: accounting for taste in students' evaluations of teaching effectiveness. J Appl Soc Psychol. 2009;39:1609-1638. http://dx.doi.org/10.1111/j.1559-1816.2009.00497.x
24. Remedios R, Lieberman DA. I liked your course because you taught me well: the influence of grades, workload, expectations and goals on students' evaluations of teaching. Br Educ Res J. 2008;34:91-115. http://dx.doi.org/10.1080/01411920701492043
25. Chen Y, Hoshower LB. Student evaluation of teaching effectiveness: an assessment of student perception and motivation. Assess Eval High Educ. 2003;28:71-88. http://dx.doi.org/10.1080/02602930301683
26. Worthington AC. The impact of student perceptions and characteristics on teaching evaluations: a case study in finance education. Assess Eval High Educ. 2002;27:49-64. http://dx.doi.org/10.1080/02602930120105054
27. Kember D, Wong A. Implications for evaluation from a study of students' perceptions of good and poor teaching. High Educ. 2000;40:69-97. http://dx.doi.org/10.1023/A:1004068500314
28. Marsh HW. The influence of student, course and instructor characteristics on evaluations of university teaching. Am Educ Res J. 1980;17:219-237. http://dx.doi.org/10.3102/00028312017002219
29. Marsh HW. Multidimensional ratings of teaching effectiveness by students from different academic settings and their relation to student/course/instructor characteristics. J Educ Psychol. 1983;75:150-166. http://dx.doi.org/10.1037/0022-0663.75.1.150
30. Marsh HW, Roche L. Effects of grading leniency and low workload on students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? J Educ Psychol. 2000;92:202-228. http://dx.doi.org/10.1037/0022-0663.92.1.202

31. Rowden GV, Carlson RE. Gender issues and students' perceptions of instructors' immediacy and evaluation of teaching and course. Psychol Rep. 1996;78:835-839. http://dx.doi.org/10.2466/pr0.1996.78.3.835
32. Goos M, Gannaway D, Hughes C. Assessment as an equity issue in higher education: comparing the perceptions of first year students, course coordinators, and academic leaders. Aust Educ Res. 2011;38:95-107. http://dx.doi.org/10.1007/s13384-010-0008-2
33. Davies M, Hirschberg J, Lye J, Johnson C, McDonald I. Systematic influences on teaching evaluations: the case for caution. Aust Econ Pap. 2007;46:18-38. http://dx.doi.org/10.1111/j.1467-8454.2007.00303.x
34. Blackhart GC, Peruche BM, DeWall CN, Joiner TE. Factors influencing teaching evaluations in higher education. Teach Psychol. 2006;33:37-39. http://dx.doi.org/10.1207/s15328023top3301_9
35. Dwinell PL, Higbee JL. Students' perceptions of the value of teaching evaluations. Percept Mot Skills. 1993;76:995-1000. http://dx.doi.org/10.2466/pms.1993.76.3.995
36. Burdsal CA, Bardo JW. Measuring student's perceptions and teaching dimensions of evaluation. Educ Psychol Meas. 1986;46:63-79. http://dx.doi.org/10.1177/0013164486461006
37. Mullen GE, Tallant-Runnels MK. Student outcomes and perceptions of instructors' demands and support in online and traditional classrooms. Internet High Educ. 2006;9:257-266. http://dx.doi.org/10.1016/j.iheduc.2006.08.005
38. Theall M, Franklin J. Looking for bias in all the wrong places: a search for truth or a witch hunt in student ratings of instruction? New Dir Inst Res. 2001;109:45-56. http://dx.doi.org/10.1002/ir.3
39. Feldman KA. The significance of circumstances for college students' ratings of their teachers and courses. Res High Educ. 1979;10:149-172. http://dx.doi.org/10.1007/BF00976227
40. Sojka J, Gupta AK, Deeter-Schmelz DR. Student and faculty perceptions of student evaluations of teaching: a study of similarities and differences. Coll Teach. 2002;50:44-49. http://dx.doi.org/10.1080/87567550209595873
41. Berk RA. Survey of 12 strategies to measure teaching effectiveness. Int J Teach Learn High Educ. 2005;17:48-62.
42. Greenwald AG, Gillmore GM. Grading leniency is a removable contaminant of student ratings. Am Psychol. 1997;52:1209-1217. http://dx.doi.org/10.1037/0003-066X.52.11.1209
43. Gigliotti RJ, Buchtel FS. Attributional bias and course evaluations. J Educ Psychol. 1990;82:341-351. http://dx.doi.org/10.1037/0022-0663.82.2.341
44. Doyle KO, Crichton LL. Student, peer and self evaluations of college instructors. J Educ Psychol. 1978;70:815-826. http://dx.doi.org/10.1037/0022-0663.70.5.815
45. Schum TR, Koss R, Yindra KJ, Nelson DB. Students' and residents' ratings of teaching effectiveness in a department of paediatrics. Teach Learn Med. 1993;5:128-132. http://dx.doi.org/10.1080/10401339309539606
46. Aleamoni LM. Student rating myths vs research from 1924-1998. J Pers Eval Educ. 1999;13:153-166.
47. Arah OA, Heineman MJ, Lombarts K. Factors influencing residents' evaluatons of clinical faculty member teaching qualities and role model status. Med Educ. 2012;46:381-389. http://dx.doi.org/10.1111/j.1365-2923.2011.04176.x
48. Tews MJ, Tracey JB. Enhancing formal interpersonal skills training through post-training supplements. Cornell Hosp Q. 2007;7:4-20.
49. Tews MJ, Tracey JB. Helping managers help themselves: the use and utility of on-the-job interventions to improve the impact of interpersonal skills training. Cornell Hosp Q. 2009;50:245-258. http://dx.doi.org/10.1177/1938965509333520
50. Langenfeld SJ, Helmer SD, Cusick TE, Smith RS. Do strong resident teachers help medical students on objective examinations of knowledge? J Surg Educ. 2011;68:350-354. http://dx.doi.org/10.1016/j.jsurg.2011.05.003
51. Schneider JR, Coyle JJ, Ryan ER, Bell RH Jr, DaRosa DA. Implementation and evaluation of a new surgical residency model. J Am Coll Surg. 2007;205:393-404. http://dx.doi.org/10.1016/j.jamcollsurg.2007.05.013
52. Kember D, Leung D. Development of a questionnaire for assessing student's perceptions of the teaching and learning environment and its use in quality assurance. Learn Environ Res. 2009;12:15-29. http://dx.doi.org/10.1007/s10984-008-9050-7
53. Scott A, Jeon SH, Joyce CM, Humphreys JS, Kalb G, Witt J, Leahy A. A randomised trial and economic evaluation of the effect of response mode on response rate, response bias, and item non-response in a survey of doctors. BMC Med Res Methodol. 2011;11:126-138. http://dx.doi.org/10.1186/1471-2288-11-126
54. Smither JW, London M, Vasilopoulos NL, Reilly RR, Millsap RE, Salvemini N. An examination of the effects of an upward feedback program over time. Pers Psychol. 1995;48:1-34. http://dx.doi.org/10.1111/j.1744-6570.1995.tb01744.x
55. Ahearn D, Bhat S, Lakinson T, Baker P. Maximising responses to quality assurance surveys. Clin Teach. 2011;8:258-262. http://dx.doi.org/10.1111/j.1743-498X.2011.00477.x
56. Antonioni D, Park H. The relationship between rater affect and three sources of 360-degree feedback ratings. J Manage. 2001;27:479-495. http://dx.doi.org/10.1177/014920630102700405
57. Tsui AS, Barry B. Research notes: interpersonal affect and rating errors. Acad Manage J. 1986;29:586-599. http://dx.doi.org/10.2307/256225
58. Albanese M. Rating educational quality: factors in the erosion of professional standards. Acad Med. 1 999;74:652-658.
59. Ryan AM, Brutus S, Greguras GJ, Hakel MD. Receptivity to assessment-based feedback for management development. J Ma-

nag Dev. 2000;19:252-276. http://dx.doi.org/10.1108/02621710010322580
60. Eva KW, Armson H, Holmboe E, Lockyer J, Loney E, Mann K, Sargeant J. Factors influencing responsivenss to feedback: on the interplay between fear, confidence, and reasoning processes. Adv Health Sci Educ Theory Pract. 2012;17:15-26. http://dx.doi.org/10.1007/s10459-011-9290-7
61. Antonioni D. The effects of feedback accountability on upward appraisal ratings. Pers Psychol. 1994;47:349-356. http://dx.doi.org/10.1111/j.1744-6570.1994.tb01728.x
62. Goodwin J, Yeo TY. Two factors affecting internal audit independence and objectivity: evidence from Singapore. Int J Audit. 2001;5:107-125. http://dx.doi.org/10.1111/j.1099-1123.2001.00329.x
63. Grava-Gubins I, Scott S. Effects of various methodologic strategies: survey response rates among Canadian physicians and physicans-in-training. Can Fam Physician. 2008;54:1424-1430.
64. Owen JP. A survey of the provision of educational supervision in occupational medicine in the Armed forces. Occup Med. 2005;55:227-233. http://dx.doi.org/10.1093/occmed/kqi030
65. Fiander A. Evaluation of flexible senior registrar training in obstetrics and gynaecology. Br J Obstet Gynaecol. 1995;102:461-466. http://dx.doi.org/10.1111/j.1471-0528.1995.tb11318.x
66. Risucci DA, Lutsky L, Rosati RJ, Tortolani AJ. Reliability and accuracy of resident evaluations of surgical faculty. Eval Health Prof. 1992;15:313-324. http://dx.doi.org/10.1177/016327879201500304
67. Roch SG, McNall LA. An investigation of factors influencing accountability and performance ratings. J Psychology. 2007;141:499-524. http://dx.doi.org/10.3200/JRLP.141.5.499-524
68. Antonioni D. Predictors of upward feedback ratings. J Manage Issues. 1999;11:26-36.
69. Bettenhausen KL, Fedor DB. Peer and upward appraisals: a comparison of their benefits and problems. Group Organ Manage. 1997;22:236-263. http://dx.doi.org/10.1177/1059601197222006
70. Westerman JW, Rosse JG. Reducing the threat of rater nonparticipation in 360-degree feedback systems: an exploratory examination of antecedents to participation in upward ratings. Group Organ Manage. 1997;22:288-309. http://dx.doi.org/10.1177/1059601197222008
71. Mathews BP, Redman T. The attitudes of service industry managers towards upward appraisals. Career Dev Int. 1997;2:46-53. http://dx.doi.org/10.1108/13620439710157498
72. Reid P, Levy G. Subordinate appraisal of managers: a useful tool for the NHS? Health Manpow Manage. 1997;23:68-72. http://dx.doi.org/10.1108/09552069710166698
73. Atwater L, Roush P, Fischthal A. The influence of upward feedback on self- and follower rating. Pers Psychol. 1995;48:35-59. http://dx.doi.org/10.1111/j.1744-6570.1995.tb01745.x
74. Redman T, McElwee G. Upward appraisal of lecturuers: lessons from industry? Educ Train. 1993;35:20-26. http://dx.doi.org/10.1108/EUM0000000000297
75. Redman T, Snape E. Upward and onward: can staff appraise their managers? Pers Rev. 1992;21:32-46. http://dx.doi.org/10.1108/00483489210021044
76. Chan D, Ip WY. Perception of hospital learning environment: a survey of Hong Kong nursing students. Nurse Educ Today. 2007;27:677-684. http://dx.doi.org/10.1016/j.nedt.2006.09.015
77. Henderson A, Beattie H, Boyde M, Storrie K, Lioyd B. An evaluation of the first year of a collaborative tertiary-industry curriculum as measured by students perceptions of their clinical learning environment. Nurse Educ Pract. 2006;6:207-213. http://dx.doi.org/10.1016/j.nepr.2006.01.002
78. Perli S, Brugnolli A. Italian nursing students' perception of their clinical learning environment as measured with the CLEI tool. Nurse Educ Today. 2009;29:886-890. http://dx.doi.org/10.1016/j.nedt.2009.05.016
79. Severinsson E, Sand A. Evaluation of the clinical supervision and professional development of student nurses. J Nurs Manag. 2010;18:669-677. http://dx.doi.org/10.1111/j.1365-2834.2010.01146.x
80. Midgley K. Pre-registration student nurses perception of the hsopital-learning environment during clinical placements. Nurse Educ Today. 2006;26:338-345. http://dx.doi.org/10.1016/j.nedt.2005.10.015
81. Cohan JA. "I didn't know" and "I was only doing my job": has corporate governance careened out of control? A case study of Enron's information myopia. J Bus Ethics. 2002;40:275-299. http://dx.doi.org/10.1023/A:1020506501398
82. Palmgren PJ, Chandratilake M. Perception of educational environment among undergraduate students in a chiropractic training institution. J Chiropract Educ. 2011;25:151-163. http://dx.doi.org/10.7899/1042-5055-25.2.151
83. Raikkonen O, Perala ML, Kahanpaa A. Staffing adequacy, supervisory support and quality of care in long-term settings: staff perceptions. J Adv Nurs. 2007;60:615-626. http://dx.doi.org/10.1111/j.1365-2648.2007.04443.x
84. Rabow M, Gargani J, Cooke M. Do as I say: curricular discordance in medical schools end-of-lief care education. J Palliat Med. 2007;10:759-769. http://dx.doi.org/10.1089/jpm.2006.0190
85. Kolarik RC, Walker G, Arnold RM. Paediatric residents education in palliative care: a needs assessment. Pediatrics. 2006;117:1949-1954. http://dx.doi.org/10.1542/peds.2005-1111
86. Smith KL, Tichenor CJ, Schroeder M. Orthopaedic residency training: a survey of the graduates' perspective. J Orthop Sports Phys Ther. 1999;29:635-651. http://dx.doi.org/10.2519/jospt.1999.29.11.635
87. Paul P, Olson J, Jackman D, Gauthier S, Gibson B, Kabotoff W, Weddell A, Hungler K. Perceptions of extrinsic factors that contribute to a nursing internship experience. Nurse Educ Today.

2011;31:763-767. http://dx.doi.org/10.1016/j.nedt.2010.11.016
88. Braine ME, Parnell J. Exploring student's perceptions and experience of personal tutors. Nurse Educ Today. 2011;31:904-910. http://dx.doi.org/10.1016/j.nedt.2011.01.005
89. Brugnolli A, Perli S, Viviani D, Saiani L. Nursing students' perceptions of tutorial strategies during clinical learning instructions. Nurse Educ Today. 2011;31:152-156. http://dx.doi.org/10.1016/j.nedt.2010.05.008
90. Heffernan C, Heffernan E, Brosnan M. Evaluating a preceptorship programme in south west Ireland: Perceptions of preceptors and undergraduate students. J Nurs Manag. 2009;17:539-549. http://dx.doi.org/10.1111/j.1365-2834.2008.00935.x
91. Kelly C. Student's perceptions of effective clinical teaching revisited. Nurse Educ Today. 2007;27:885-892. http://dx.doi.org/10.1016/j.nedt.2006.12.005
92. Ranse K, Grealish L. Nursing students' perceptions of learning in the clinical setting of the dedicated education unit. J Adv Nurs. 2007;58:171-179. http://dx.doi.org/10.1111/j.1365-2648.2007.04220.x
93. Beecroft PC, Santner S, Lacy ML, Kunzman L, Dorey F. New graduate nurses' perceptions of mentoring: six-year programme evaluation. J Adv Nurs. 2006;55:736-747. http://dx.doi.org/10.1111/j.1365-2648.2006.03964.x
94. Sit JW, Chung JW, Chow MC, Wong T. Experiences of online learning: students' perspective. Nurse Educ Today. 2005;25:140-147. http://dx.doi.org/10.1016/j.nedt.2004.11.004.
95. O'Connor K, Joshi N, Rasburn N, Molyneux M. Thoracic anaesthesia training: the national 'One Lung' survey. Anaesthesia. 2011;66:325-326. http://dx.doi.org/10.1111/j.1365-2044.2011.06676.x
96. Luks AM, Smith CS, Robins L, Wipf JE. Resident perceptions of the educational value of night float rotations. Teach Learn Med. 2010;22:196-201. http://dx.doi.org/10.1080/10401334.2010.488203
97. Turnball C, Baker P, Allen S. A comparison of three different quality assurance systems for higher medical training. Clin Med. 2007;7:486-491.
98. Biller CK, Antonacci AC, Pelletier S, Homel P, Spann C, Cunningham MJ, Eavey RD. The 80-hour work guidelines and resident survey perceptions of quality. J Surg Res. 2006;135:275-281. http://dx.doi.org/10.1016/j.jss.2006.04.010
99. Carpenter RO, Spooner J, Arbogast PG, Tarplay JL, Griffin MR, Lomis KD. Work-hour restrictions as an ethical dilemma for residents. Am J Surg. 2006;191:527-532. http://dx.doi.org/10.1016/j.cursur.2006.06.003
100. Brasher AE CS, Hauge LS, Prinz RA, Neumayer LA, Baker CC, Soybel DI, Freischlag JA, Jeekel JH. Medical students' perceptions of resident teaching: have duty hours regulations had an impact? Ann Surg. 2005;242:548-555. http://dx.doi.org/10.1097/01.sla.0000184192.74000.6a
101. Busari JO, Wegglaar NM, Knottnerus AC, Greidanus PM, Scherpbier AJ. How medical residents perceieve the quality of supervision provided by attending doctors in the clinical setting. Med Educ. 2005;39:696-703. http://dx.doi.org/10.1111/j.1365-2929.2005.02190.x
102. Ansari WE, Oskrochi R. What 'really' affects health professions students' satisfaction with their educational experience? Implications for practice and research. Nurse Educ Today. 2004;24:644-655. http://dx.doi.org/10.1016/j.nedt.2004.09.002
103. Basu CB, Chen LM, Hollier LH Jr, Shenaq SM. The effect of the accreditation council for graduate medical education duty hours policy on plastic surgery resident education and patient care: an outcomes study. Plast Reconstr Surg. 2004;114:1878-1886. http://dx.doi.org/10.1097/01.PRS.0000142768.07468.64
104. Whang EE, Mello MM, Ashley SW, Zinner MJ. Implementing resident work hour limitations: lessons from the New York State experience. Ann Surg. 2003;237:449-455. http://dx.doi.org/10.1097/01.SLA.0000059966.07463.19
105. Devlin MF, McCaul JA, Currie WJ. Trainees perceptions of UK Maxillofacial training. Br J Oral Maxillofac Surg. 2002;40:424-428. http://dx.doi.org/10.1016/S0266-4356(02)00200-0
106. Metcalfe DH, Matharu M. Students' perception of good and bad teaching: report of a critical incident study. Med Educ. 1995;29:193-197. http://dx.doi.org/10.1111/j.1365-2923.1995.tb02829.x
107. Barrett E, Barry H, Guruswamy S, McCarthy M, Kavanagh E. What trainees really think: the 2009 and 2010 national trainee survey of trainees's perception of their training in Ireland. In: 20th European Congress of Psychiatry; 2012 Mar 3-6; Prague, Czech Republic.
108. Steiner IP, Yoon PW, Kelly KD, Diner BM, Blitz S, Donoff MG, Rowe BH. The influence of residents training level on their evaluation of clinical teaching faculty. Teach Learn Med. 2005;17:42-48. http://dx.doi.org/10.1207/s15328015tlm1701_8
109. Getz TA, Evens RG. Residencies in diagnostic radiology and perception of residents: 1987 A3CR2 survey. Invest Radiol. 1988;23:308-311. http://dx.doi.org/10.1097/00004424-198804000-00012
110. Berber M. How can faculty course survey be made more meaningful? Surv Land Inf Sci. 2011;71:13-19.
111. Girard DE, Choi D, Dickey J, Dickerson D, Bloom JD. A comparison study of career satisfication and emotional states between primary care and speciality residents. Med Educ. 2006;49:79-86. http://dx.doi.org/10.1111/j.1365-2929.2005.02350.x
112. Antiel R, Van Arendonk K, Reed D, Terhune JP, Tarpley JL, Porterfield JR, Hall DE, Joyce DL, Wightman SC, Horvath KD, Heller SF, Farley DR. Surgical training, duty-hour restrictions, and implications for meeting the accreditations council for graduate medical education core competencies: views of surgical interns compared with program directors. Arch Surg. 2012;147:536-541. http://dx.doi.org/10.1001/archsurg.2012.89

113. Lin GA, Beck DC, Stewart AL, Garbutt JM. Resident perceptions of the impact of work hour limitations. J Gen Intern Med. 2007;22:969-975. http://dx.doi.org/10.1007/s11606-007-0223-3
114. Ratanawongsa N, Bolen S, Howell EE, Kern D, Sisson S, Larriviere D. Residents' perceptions of professionalism in training and practice: barriers, promoters, and duty hour requirements. J Gen Intern Med. 2006;21:758-763. http://dx.doi.org/10.1111/j.1525-1497.2006.00496.x
115. Thangaratinam S, Yanamandra SR, Deb S, Coomarasamy A. Specialist training in obstetrics and gynaecology: a survey on work-life balance and stress among trainees in UK. J Obstet Gynaecol. 2006;26:302-304. http://dx.doi.org/10.1080/01443610600594773
116. Kanashiro J, McAleer S, Roff S. Assessing the educational environment in the operating room: a measure of resident perception at one Canadian institution. Surgery. 2006;139:150-158. http://dx.doi.org/10.1016/j.surg.2005.07.005
117. Blue AV, Griffith CH, Wilson J, Sloan DA, Schwartz RW. Surgical teaching quality makes a difference. Am J Surg. 1999;177:86-89. http://dx.doi.org/10.1016/S0002-9610(98)00304-3
118. Watling C, Driessen E, Van der Vleuten C, Lingard L. Learning from clinical work: the roles of learning cues and credibility judgements. Med Educ. 2012;46:192-200. http://dx.doi.org/10.1111/j.1365-2923.2011.04126.x
119. Iqbal M, Khizar B. Medical Students' perceptions of teaching evaluations. Clin Teach. 2009;6:69-72. http://dx.doi.org/10.1111/j.1743-498X.2009.00268.x
120. Watling C, Kenyon CF, Zibrowski EM, Schulz V, Goldszmidt MA, Singh I, Maddocks HL, Lingard L. Rules of engagement: residents' perceptions of the in-training evaluation process. Acad Med. 2008;83:S97-S100. http://dx.doi.org/10.1097/ACM.0b013e318183e78c
121. Cannon G, Keitz S, Holland G, Chang B, Byrne J, Tomolo A, Aron DC, Wicker AB, Kashner TM. Factors determining medical students' and residents' satisfaction during VA-based training: Findings from the VA learners' perceptions survey. Acad Med. 2008;83:611-620. http://dx.doi.org/10.1097/ACM.0b013e3181722e97
122. Pearce I, Royle J, O'Flynn K, Payne S. The record of in-training assessments (RITAs) in urology: an evaluation of trainee perceptions. Ann R Coll Surg Engl. 2003;85:351-354. http://dx.doi.org/10.1308/003588403769162495
123. Conigliaro J, Frishman WH, Lazar EJ, Creons L. Internal medicine housestaff and attending physician perceptions of the impact of the New York State Section 405 regulations on working conditions and supervision of residents in two training programs. J Gen Intern Med. 1993;8:502-507. http://dx.doi.org/10.1007/BF02600112
124. Dech B, Abikoff H, Koplewicz HS. A survey of child and adolescent psychiatry residents: perceptions of the ideal training program. J Am Acad Child Adolesc Psychiatry. 1990;29:946-949. http://dx.doi.org/10.1097/00004583-199011000-00019
125. Yarris L, Linden J, Hern G, Lefebvre C, Nestler DM, Fu R, Choo E, LaMantia J, Burnett P; Emergency Medicine Education Research Group. Attending and resident satisfaction with feedback in the emergency department. Acad Emerg Med. 2009;16:S76-S78. http://dx.doi.org/10.1111/j.1553-2712.2009.00592.x
126. Sargeant J, Mann K, Sinclair D, Van der vleuten C, Metsemakers J. Understanding the influence of emotions and reflection upon multi-source feedback acceptance and use. Adv Health Sci Educ Theory Pract. 2008;13:275-288. http://dx.doi.org/10.1007/s10459-006-9039-x
127. Solomon DJ, Speer AJ, Rosebraugh CJ, DiPette DJ. The reliability of medical student ratings of clinical teaching. Eval Health Prof. 1997;20:343-352. http://dx.doi.org/10.1177/016327879702000306
128. Sender Liberman A, Liberman M, Steinert Y, Mcleod P, Meterissian S. Surgery residents and attending surgeons have different perceptions of feedback. Med Teach. 2005;27:470-472. http://dx.doi.org/10.1080/01425905001291183
129. Johnson NR, Chen J. Medical student evaluation of teaching quality between obstetrics and gynaecology residents and faculty as clinical preceptors in ambulatory gynaecology. Am J Obstet Gynecol. 2006;195:1479-1483. http://dx.doi.org/10.1016/j.ajog.2006.05.038
130. Windish DM, Knight AM, Wright SM. Clinician-teachers' self assessments versus learners' perceptions. J Gen Intern Med. 2004;19:554-557. http://dx.doi.org/10.1111/j.1525-1497.2004.30014.x
131. Tortolani A, Risucci DA, Rosati RJ. Resident evaluation of surgical faculty. J Surg Res. 1991;51:186-191. http://dx.doi.org/10.1016/0022-4804(91)90092-Z
132. O'Brien M, Brown J, Ryland I, Shaw N, Chapman T, Gillies R, Graham D. Exploring the views of second-year Foundation Programme doctors and their educational supervisors during a deanary-wide pilot Foundation Programme. Postgrad Med J. 2006;82:813-816. http://dx.doi.org/10.1136/pgmj.2006.049676
133. Claridge J, Forrest Calland J, Chandrasekhara V, Young JS, Sanfey H, Schirmer BD. Comparing resident measurements to attending surgeons self-perceptions of surgical educators. Am J Surg. 2003;185:323-327. http://dx.doi.org/10.1016/S0002-9610(02)01421-6
134. Robbins TL, DeNisi AS. A closer look at interpersonal affect as a distinct influence on coginitive processing in performance evaluations. J Appl Psychol. 1994;79:341-353. http://dx.doi.org/10.1037/0021-9010.79.3.341
135. Ramsey PG, Gillmore GM, Irby DM. Evaluating clinical teaching in medicine clerkship: relationship of instructor experience and training setting to ratings of tutor effectiveness. J Gen Intern Med. 1988;3:351-355.

136. Hayward RA, Williams BC, Gruppen LD, Rosenbaum D. Measuring attending physician performance in a general medicine outpatient clinic. J Gen Intern Med. 1995;10:504-510. http://dx.doi.org/10.1007/BF02602402
137. Sargeant J, Mann K, Suzanne F. Exploring family physicians' reactions to multisource feedback: perception of credibility and usefulness. Med Educ. 2005;39:497-504. http://dx.doi.org/10.1111/j.1365-2929.2005.02124.x
138. Brett JF, Atwater LE. 360 feedback: accuracy, reactions, and perceptions of usefulness. J Appl Psychol. 2001;86:930-942. http://dx.doi.org/10.1037/0021-9010.86.5.930
139. Barclay LJ, Skarlicki DP, Pugh SD. Exploring the role of emotions in injustice perceptions and retaliation. J Appl Psychol. 2005;90:629-643. http://dx.doi.org/10.1037/0021-9010.90.4.629
140. Irby DM, Gilmore GM, Ramsey PG. Factors affecting ratings of clinical teachers by medical students and residents. Med Educ. 1987;62:1-7. http://dx.doi.org/10.1097/00001888-198701000-00001
141. Paice E, Aitken M, Houghton A, Firth-Cozens J. Bullying among doctors in training: cross sectional questionnaire survey. Br Med J. 2004;324:658-659. http://dx.doi.org/10.1136/bmj.38133.502569.AE
142. Ryland I, Brown J, O'Brien M, Graham D, Gillies R, Chapman T, Shaw N. The portfolio: how was it for you? Views of F2 doctors from the Mersey Deanery Foundation Pilot. Clin Med. 2006;6:378-380.
143. Watling C, Lingard L. Toward meaningful evaluation of medical trainees: The influence of participants' perceptions of the process. Adv Health Sci Educ Theory Pract. 2010;17:183-194. http://dx.doi.org/10.1007/s10459-010-9223-x
144. Tochel C, Haig A, Hesketh A, Cadzow A, Beggs K, Colthart I, Peacock H. The effectiveness of portfolios for post-graduate assessment and education: BEME Guide No.12. Med Teach. 2009;31:299-318. http://dx.doi.org/10.1080/01421590902883056
145. Rose JS, Waibel BH, Schenarts PJ. Disparity between resident and faculty surgeons' perceptions of preoperative preparation, intraoperative teaching, and postoperative feedback. J Surg Educ. 2011;68:459-464. http://dx.doi.org/10.1016/j.jsurg.2011.04.003
146. Govaerts M, Van Der Vleuten C, Schuwirth L, Muijtjens A. The use of observational diaries in in-training evaluation: student perceptions. Adv Health Sci Educ Theory Pract. 2005;10:171-188. http://dx.doi.org/10.1007/s10459-005-0398-5
147. Williams BC, Pillsbury MS, Stern DT, Grum CM. Comparison of resident and medical student evaluation of faculty teaching. Eval Health Prof. 2001;24:53-60. http://dx.doi.org/10.1177/016327801220347 86
148. Tourish D, Robson P. Critical upward feedback in organisations: processes, problems and implications for commmunication management. J Comm Manag. 2003;8:150-167. http://dx.doi.org/10.1108/13632540410807628
149. Surratt CK, Desselle SP. Pharmacy students' perceptions of a teaching evaluation process. Am J Pharm Educ. 2007;71:6.
150. Ilgen DR, Fisher CD, Taylor MS. Consequences of individual feedback on behavior in organizations. J Appl Psychol. 1979;64:349-371. http://dx.doi.org/10.1037/0021-9010.64.4.349
151. Bing-You RG, Paterson J, Mark AL. Feedback falling on deaf ears: residents' receptivity to feedback tempered by sender credibility. Med Teach. 1997;19:40-44. http://dx.doi.org/10.3109/01421599709019346
152. Fallon SM, Creon LG, Shelov SP. Teachers' and students' ratings of clinical teaching and teachers' opinions on use of student evaluations. Med Educ. 1987;62:435-438.
153. Stritter FT, Hain JD, Grimes DA. Clinical teaching re-examined. J Med Educ. 1975;62:1-7.
154. Shellenberger S, Mahan JM. A factor analytic study of teaching in off-campus general practice clerkships. Med Educ. 1982;16:151-155. http://dx.doi.org/10.1111/j.1365-2923.1982.tb01076.x
155. Cohen R, MacRae H, Jamieson C. Teaching effectiveness of surgeons. Am J Surg. 1996;171:612-614. http://dx.doi.org/10.1016/S0002-9610(97)89605-5
156. Dolmans D, Van Luijk SJ, Wolfhagen I, Scherpbier A. The relationship between professional behaviour grades and tutor performance ratings in problem-based learning. Med Educ. 2006;40:180-186. http://dx.doi.org/10.1111/j.1365-2929.2005.02373.x
157. Donnelly M, Wooliscroft J. Evaluation of clinical instructors by third-year medical students. Acad Med. 1989;64:159-164. http://dx.doi.org/10.1097/00001888-198903000-00011
158. Irby D, Rakestraw P. Evaluating clinical teaching in medicine. Med Educ. 1981;56:181-186.
159. Parikh A, Mcreelis K, Hodges B. Student feedback in problem based learning: a survey of 103 final year students accross five Ontario medical schools. Med Educ. 2001;35:632-663. http://dx.doi.org/10.1046/j.1365-2923.2001.00994.x
160. Wilson FC. Teaching by residents. Clin Orthop Relat Res. 2007;454:247-250. http://dx.doi.org/10.1097/BLO.0b013e31802b4944
161. De SK, Henke PK, Ailawadi G, Dimick JB, Colletti LM. Attending, house officer, and medical student perceptions about teaching in the third-year medical student perceptions about teaching in the third-year medical school general surgery clerkship. J Am Coll Surg. 2004;199:932-942. http://dx.doi.org/10.1016/j.jamcollsurg.2004.08.025
162. Duffield KE, Spencer JA. A survey of medical students' views about the purposes and fairness of assessment. Med Educ. 2002;36:879-886. http://dx.doi.org/10.1046/j.1365-2923.2002.01291.x
163. Tiberius RG, Sackin HD, Slingerland JM, Jubas K, Bell M, Matlow A. The influence of student evaluative feedback on the improvement of clinical teaching. J High Educ. 1989;60:665-681.
164. Gil DH, Heins DM, Jones PB. Perceptions of medical school faculty members and students on clinical clerkship feedback. Med Educ. 1984;59:856-864. http://dx.doi.org/10.1097/0000

1888-198411000-00003
165. Pfeifer MP, Peterson HR. The influence of student interest on teaching evaluation. J Gen Intern Med. 1991;6:141-144. http://dx.doi.org/10.1007/BF02598312
166. Cardy RL, Dobbins GH. Affect and appraisal accuracy: liking as an integral dimension in evaluating performance. J Appl Psychol. 1986;71:672-678. http://psycnet.apa.org/doi/10.1037/0021-9010.71.4.672
167. Henzi D, Davis E, Jasinevicius R, Hendricson W, Clintron L, Isaacs M. Appraisal of the dental school learning environment: the students' view. J Dental Educ. 2005;69:1137-1147.
168. Parker T, Carlisle C. Project 2000 students' perceptions of their training. J Adv Nurs. 1996;24:771-778. http://dx.doi.org/10.1046/j.1365-2648.1996.25416.x
169. Cooke M, Mitchell M, Moyle W. Application and student evaluation of a clinical progression portfolio: a pilot. Nurse Educ Pract. 2010;10:227-232. http://dx.doi.org/10.1016/j.nepr.2009.11.010
170. Myall M, Levett-Jones T, Lathlean J. Mentorship in contemporary practice: the experiences of nursing students and practice mentors. J Clin Nurs. 2008;17:1834-1842. http://dx.doi.org/10.1111/j.1365-2702.2007.02233.x
171. Kjaer N, Maagaard R, Wied S. Using an online portfolio in postgraduate training. Med Teach. 2006;28:708-712. http://dx.doi.org/10.1080/01421590601047672.
172. Hrisos S, Illing J, Burk J. Portfolio learning for foundation doctors: early feedback on its use in the clinical workplace. Med Educ. 2008;42:214-223. http://dx.doi.org/10.1111/j.1365-2923.2007.02960.x
173. Beckman M, Lee M, Mandrekar J. A comparison of clinical teaching evaluations by resident and peer physicians. Med Teach. 2004;26:321-325. http://dx.doi.org/10.1080/01421590410001678984
174. Mattern WD, Weinholtz D, Friedman CP. The attending physician as a teacher. N Engl J Med. 1983;308:1129-1132. http://dx.doi.org/10.1056/NEJM198305123081904
175. Kendrick SB, Simmons J, Richards B, L R. Resident's perception of their teachers; facilitative behaviour and learning value of rotations. Med Educ. 1993;27:55-61. http://dx.doi.org/10.1111/j.1365-2923.1993.tb00229.x
176. Keitz S, Gilman S, Breen A, Graber M. Measuring the quality of veterans affairs (VA) clinical training: a learners perception survey of medical residents in VA medical centres. J Gen Intern Med. 2002;17:228.
177. Moalem J, Salzman P, Ruan DT, Cherr GS, Freiburg CB, Farkas RL, Brewster L, James TA. Should all duty hours be the same? Results of a national survey of surgical trainees. J Am Coll Surg. 2009;209:47-54, 54.e1-2. http://dx.doi.org/10.1016/j.jamcollsurg.2009.02.053
178. Sargeant J, McNaughton E, Mercer S, Murphy D, Sullivan P, Bruce DA. Providing feedback: exploring a model (emotion, content, outcomes) for facilitating multisource feedback. Med Teach. 2011;33:744-749. http://dx.doi.org/10.3109/0142159X.2011.577287
179. Schuh LA, Khan MA, Harle H, Southerland AM, Hicks WJ, Falchook A, Schultz L, Finney GR. Pilot trial of IOM duty hour recommendations in neurology residency programs: unintended consequences. Neurology. 2011;77:883-887. http://dx.doi.org/10.1212/WNL.0b013e31822c61c3
180. Vasudev A, Vasudev K, Thakkar P. Trainees' perception of the Annual Review of Competence Progression: 2-year survey. Psychiatrist. 2010;34:396-399. http://dx.doi.org/10.1192/pb.bp.109.028522
181. Ellrodt AG. Introduction of total quality management (TQM) into an internal medicine residency. Acad Med. 1993;68:817-823. http://dx.doi.org/10.1097/00001888-199311000-00002
182. Harrison R, Allen E. Teaching internal medicine residents in the new era: inpatient attending with duty-hour regulations. J Gen Intern Med. 2006;21:447-452. http://dx.doi.org/10.1111/j.1525-1497.2006.00425.x
183. Dola C, Nelson L, Lauterbach J, Degefu S, Pridjian G. Eighty hour work reform: faculty and resident perceptions. Am J Obstet Gynecol. 2006;195:1450-1456. http://dx.doi.org/10.1016/j.ajog.2006.06.074
184. Cohn DE, Roney JD, O'Malley DM, Valmadre S. Residents' perspectives on surgical training and the resident-fellow relationship: comparing residency programs with and without gynecological oncology fellowships. Int J Gynecol Cancer. 2008;18:199-204. http://dx.doi.org/10.1111/j.1525-1438.2007.00986.x
185. Fisher VL, Barnes Y, Olson EA, Sheens MA, Nieder ML. Mid-level practitioner-physician collaboration in pediatric HSCT programs. Biol Blood Marrow Transplant. 2010;16:S329. http://dx.doi.org/10.1016/j.bbmt.2009.12.521
186. Pankhania M, Ghouri A, Sahota RS, Carr E, Ali K, Pau H. Special senses: changing the face of undergraduate ENT teaching. In: 6th Meeting of the South West ENT Academic Meeting; 2011 Jun; Bath, UK.
187. Welch J, Bridge C, Firth D, Forrest A. Improving psychiatry training in the Foundation Programme. Psychiatrist. 2011;35:389-393. http://dx.doi.org/10.1192/pb.bp.111.034009
188. Greysen SR, Schiliro D, Horwitz LI, Curry L, Bradley EH,. "Out of sight, out of mind": Housestaff perceptions of quality-limiting factors in discharge care at teaching hospitals. J Hosp Med. 2012;7:376-381. http://dx.doi.org/10.1002/jhm.1928
189. Mailloux C. The extent to which students' perceptions of faculties' teaching strategies, students' context, and perceptions of learner empowerment predict perceptions of autonomy in BSN students. Nurse Educ Today. 2006;26:578-585. http://dx.doi.org/10.1016/j.nedt.2006.01.013
190. Buschbacher R, Braddom RL. Resident versus program director perceptions about PM&R research training. Am J Phys Med Rehabil. 1995;74:90-100.

191. Cooke L, Hutchinson M. Doctors' professional values: results from a cohort study of United Kingdom medical graduates. Med Educ. 2001;35:735-742. http://dx.doi.org/10.1046/j.1365-2923.2001.01011.x
192. Holland RC, Hoysal N, Gilmore A, Acquilla S. Quality of training in public health in the UK: results of the first national training audit. Public Health. 2006;120:237-248. http://dx.doi.org/10.1016/j.puhe.2005.08.019
193. Sabey A, Harris M. Training in hospitals: what do GP specialist trainees think of workplace-based assessments? Educ Prim Care. 2011;22:90-9.
194. Nettleton S, Burrows R, Watt I. Regulating medical bodies? The consequences of the 'modernisation' of the NHS and the disemboiment of clinical knowledge. Sociol Health Illn. 2008;30:333-348. http://dx.doi.org/10.1111/j.1467-9566.2007.01057.x
195. Chamberlain JE, Nisker JA. Residents' attitudes to training in ethics in Canadian obstetrics and gynecology programs. Obstet Gynecol. 1995;85:783-786. http://dx.doi.org/10.1016/0029-7844(95)00019-N
196. Verhulst SJ, Distlehorst LH. Examination of nonresponse bias in a major residency follow-up study. Acad Med. 1993;68(2 Suppl):S61-S63. http://psycnet.apa.org/doi/10.1097/00001888-199302000-00033
197. Guyatt GH, Cook DJ, King D, Norman GR, Kane SL, Van Ineveld C. Effect of the framing of questionnaire items regarding satisfaction with training on residents' responses. Acad Med. 1999;74:192-194.
198. Barclay S, Todd C, Finlay I, Grande G, Wyatt P. Not another questionnaire! Maximising the response rate, predicting non-response and assessing no-response bias in postal questionnaire studies of GPs. Fam Pract. 2002;19:105-111. http://dx.doi.org/10.1093/fampra/19.1.105
199. Dipboye RL, De Pontbriand R. Correlates of employee reactions to performance appraisals and appraisal systems. J Appl Psychol. 1981;66:248-251. http://dx.doi.org/10.1037/0021-9010.66.2.248
200. Copp G, Caldwell K, Atwal A. Preparation for cancer care: perceptions of newly qualified health care professionals. Eur J Oncol Nurs. 2007;11:159-167. http://dx.doi.org/10.1016/j.ejon.2006.09.004
201. Bratt MM, Felzer HM. Perceptions of professional practice and work environment of new graduates in a nurse residency program. J Contin Educ Nurs. 2011;42:559-568. http://dx.doi.org/10.3928/00220124-20110516-03
202. Smither JW, Walker AG. Are the characteristics of narrative comments related to improvement in mutlirater feedback ratings over time? J Appl Psychol. 2004;89:575-581. http://dx.doi.org/10.1037/0021-9010.89.3.575
203. Becker J, Ayman R, Korabik K. Discrepancies in self/subordinates' perceptions of leadership behavior: leader's gender, organizational context, and leader's self-monitoring. Group Organ Manage. 2002;27:226-244. http://dx.doi.org/10.1177/1050110 2027002004
204. Mcleod PJ, James CA, Abrahamowicz M. Clinical tutor evaluation: a 5-year study by students on an in-patient service and residents in an ambulatory care clinic. Med Educ. 1993;27:48-54. http://dx.doi.org/10.1111/j.1365-2923.1993.tb00228.x
205. Bennett H, Gatrell J, Packham R. Medical appraisal: collecting evidence of performance through 360 degree feedback. Clin Manag. 2004;12:165-171.
206. Henzi D, Jasinevicius R, Hendricson W. In the students' own words: what are the strengths and weaknesses of the dental school curriculum? J Dent Educ. 2007;71:632-645.
207. Henzi D, Davis E, Jasinevicius R, Hendricson W. North American dental students' perspectives about their clinical education. J Dent Educ. 2006;70:361-377.
208. Baruch Y, Holtom B. Survey response rate levels and trends in organizational research. Hum Relat. 2008;61:1139-1160. http://dx.doi.org/10.1177/0018726708094863