# Influence of the FDI criteria in the restorations' evaluation and treatment decision in primary molars

Influência dos critérios da FDI na avaliação de restaurações e decisão de tratamento em molares decíduos

Djessica PEDROTTI[1] , Cleber Paradzinski CAVALHEIRO[1] , Rachel de Oliveira ROCHA[2] , Mariana Minatel BRAGA[3] ,
Thiago Machado ARDENGHI[2] , Tathiane Larissa LENZI[1,4] 

1 - Universidade Federal de Santa Maria, Graduate Program in Dental Science. Santa Maria, RS, Brazil.
2 - Universidade Federal de Santa Maria, Department of Stomatology. Santa Maria, RS, Brazil.
3 - Universidade de São Paulo, Departament of Orthodontics and Pediatric Dentistry. São Paulo, SP, Brazil.
4 - Universidade Federal do Rio Grande do Sul, School of Dentistry, Post-Graduate Program in Dentistry. Porto Alegre, RS, Brazil.

## ABSTRACT

**Objective:** This study investigated the impact of FDI criteria for evaluating restorations on examiners' decision-making compared with their previous personal judgment in primary teeth. Secondly, the possible factors related to changes when using the criteria, including the examiners' experience were explored. **Material and Methods:** A cross-sectional study in a dental office setting was conducted selecting 27 resin composite restorations placed in primary molars in 11 children. Examinations of the restorations were performed by five undergraduate and five graduate dental students. First, the evaluations were performed based on personal judgment, and 2 weeks later, with FDI criteria. All examiners underwent training to use the FDI criteria after the first evaluation. The consensus of two benchmark examiners was considered to be the reference standard. Initially, a descriptive analysis was performed. Multiple Poisson regressions analyses were used to identify possible associated factors with outcomes - to be less or more invasive based on the FDI criteria than personal judgment. **Results:** The use of the FDI criteria changed the examiners' decisions in approximately 15% of the cases. Irrespective of examiners' experience, there was a trend of false results (compared to the reference examiners) when a change in the treatment decision was registered by using the FDI criteria. Examiners chose a less invasive option when assessing multi-surface restorations with FDI criteria (PR=2.04, 95%CI=1.03-4.05; p=0.04). Examiners who spent more time for evaluation with FDI criteria were more invasive (PR=1.001, 95%CI=1.0001-1.002; p=0.03). Students were more invasive with the FDI criteria when examined children with higher *dmf-t* (PR=1.16, 95%CI=1.01-1.32; p=0.03). **Conclusion:** The use of the FDI criteria negatively influenced the restorations' evaluation and treatment decision in primary molars by undergraduate and graduate students.

## KEYWORDS

Clinical decision-making; Tooth, deciduous; Dental restoration failure; Composite resins; Pediatric dentistry

## RESUMO

**Objetivo:** Este estudo investigou o impacto dos critérios da FDI para avaliar restaurações na tomada de decisão feita pelos examinadores em comparação ao seu julgamento pessoal prévio em dentes decíduos. Secundariamente, foram explorados os possíveis fatores relacionados com mudanças causadas pelo uso dos critérios, incluindo a experiência dos examinadores. **Material e Métodos:** Um estudo transversal em ambiente de consultório odontológico foi conduzido selecionando 27 restaurações de resina composta realizado em molares decíduos em 11 crianças. As avaliações das restaurações foram realizadas por cinco estudantes de Graduação e cinco de Pós-graduação em Odontologia. As avaliações foram realizadas com base no julgamento pessoal e, duas semanas depois, com os critérios da FDI. Todos os examinadores foram treinados para utilizar os critérios da FDI após a primeira avaliação. O consenso de dois examinadores "padrão-ouro" foi considerado o padrão de referência. Inicialmente, uma análise descritiva foi realizada. Análises de regressão múltipla de Poisson foram

utilizadas para identificar possíveis fatores associados com os desfechos – ser mais ou menos invasivo com o uso dos critérios da FDI em relação ao julgamento pessoal. **Resultados:** A utilização dos critérios do FDI alterou as decisões dos examinadores em aproximadamente 15% dos casos. Independente da experiência dos examinadores, houve uma tendência de resultados falsos (em comparação com os examinadores de referência) quando uma mudança na decisão de tratamento foi registrada usando os critérios da FDI. Os examinadores escolheram uma opção menos invasiva quando avaliaram restaurações envolvendo múltiplas superfícies com os critérios da FDI (RP = 2,04, IC 95% = 1,03-4,05; p = 0,04). Os examinadores que levaram mais tempo para avaliação das restaurações foram mais invasivos com o uso dos critérios da FDI (RP = 1,001, IC 95% = 1,0001-1,0002; p = 0,03). Os alunos foram mais invasivos com o uso dos critérios da FDI quando examinaram crianças com maior experiência de cárie (RP = 1,16, IC 95% = 1,01-1,32; p = 0,03). **Conclusão:** O uso dos critérios da FDI influenciou negativamente a avaliação das restaurações e decisão de tratamento em dentes decíduos realizada por estudantes de graduação e pós-graduação.

## PALAVRAS-CHAVE:

Tomada de decisão clínica; Dente decíduo, Falha de restauração dentária; Resinas compostas; Odontopediatria

## INTRODUCTION

Restoration replacement is the most common treatment for managing defective restorations and is well-accepted within pediatric dentistry [1]. Recurrent caries and fractures are the main reasons for restorations' failure in posterior teeth [2,3]. However, parameters for evaluating the restorations' quality are often subjective, and slight alterations may determine the replacement, as each one has its own concept of defective restorations [4].

Most dentists are not conservative when they revisit a restoration that they performed, irrespective of the type of failure or number of surfaces [5]. Nonetheless, dentists who had placed the original restoration are significantly more likely to repair defective restorations placed in molars [5,6]. The examiners' experience also may be a determinant factor in the clinical evaluation of restorations and subsequent decision-making process. Dentists with little experience tend to perform more invasive treatments in primary teeth, or in some cases, even intervene in teeth that do not need any treatment [7]. On the other hand, clinical diagnosis may be subjective even among experienced clinicians, depending on "professional profile", some being more 'reactive' (do not act until the problem occurs) and others being more 'proactive' (interfere before the problem occurs) [8].

Therefore, the criteria proposed for standardizing the evaluation of restorative materials in clinical trials could be useful for assessing the restorations' quality placed by clinicians in their own practices [8,9]. Dental students should also be trained to use them as part of clinical evaluations to determine whether a restoration can be maintained or whether it needs repair or replacement [9].

The criteria proposed by the World Dental Federation (FDI) are based on the evaluation of biological, esthetic, and functional properties, and they have been considered more sensitive to identify changes in adhesive restorations [9,10]. Interestingly, the FDI criteria classify non-acceptable restorations in two categories: whether the restoration can be repaired and whether it must be replaced completely [9].

A recent study [11] evaluated the influence of using FDI and Caries Associated with Restorations and Sealants (CARS) criteria to assess carious lesions around restorations on the decision to replace restorations in primary molars. This decision was related to the evaluation method and not only by patients' risk factors. FDI criteria suggested more invasive treatments when compared with the CARS method [11]. Although the use of the FDI criteria by experienced examiners has been analyzed, to the best of our knowledge, no previous study evaluated if the use of FDI criteria affects the general examiners' decision-making for restoration re-intervention, reducing or increasing the overtreatment, in comparison with their previous personal judgment (without the use of standardized criteria). Thus, this study aimed to investigate the impact of the FDI criteria for evaluating restorations on examiners' decision-making compared with their previous personal judgment in primary teeth. Secondly, the possible factors related to changes

when using the criteria, including the examiners' experience were explored.

## MATERIAL AND METHODS

### Ethical concern and sample selection

The Local Research Board approved the research protocol (CAAE: 77219817.9.0000.5346) and the parents or guardians provided written informed consent. The personal information of the children was kept confidential.

A consecutive sample of children who sought treatment in the local dental clinic was used in this study. The inclusion criteria were children with primary or mixed dentition that presented at least one occlusal or occluso-proximal resin composite restoration in primary teeth. The exclusion criteria were children whose parents did not agree to participate in the study and who had restored teeth presenting signs or symptoms of pulp involvement (fistula, abscess, pulp exposure, history of spontaneous dental pain) or mobility. Children presenting these conditions in one or more teeth, but also presenting at least one eligible tooth fitting the inclusion criteria were still included in the study. An examiner (D.P.) who was not involved in the evaluations performed the sample inclusion process and visually selected those restorations that represent examples of different aspects (both those related to no need of intervention and to any type of intervention needed). The selection process was based on esthetic (color, anatomic form, staining, and lustre), functional (fracture and marginal adaptation) and biological (recurrent caries) parameters [9]. Thus, 27 resin composite restorations of 11 children, aged 5 to 7 years old, were included.

### Dental examinations

Ten examiners with different levels of clinical experience were invited to perform the assessments: five undergraduate dental students and five graduate students. The undergraduate students were in their last year of study in a Dentistry degree. The graduate students were enrolled in the master's program at the same university and had at least 2 years of clinical experience. Graduate students had completed the undergraduate course at the same school. Thus, all students had the same academic background about the parameters for evaluating the restorations' quality. Furthermore, none student had previous experience in using FDI criteria.

First, the examinations were performed based on personal judgement, and on a second occasion, 2 weeks later, the restorations were assessed according to the FDI criteria [9]. For both evaluations, examiners were guided to assess the restorations based on esthetic (color, anatomic form, staining, and lustre), functional (fracture and marginal adaptation) and biological (recurrent caries) parameters. Each criterion of the FDI can be expressed with five scores; three for acceptable (1. clinically very good; 2. clinically good; 3. clinically sufficient/satisfactory) and two for non-acceptable (4. clinically unsatisfactory – repairable restoration; 5. clinically poor – restoration replacement). For clinical decision-making, the worst grading among all parameters of the criteria was considered and then the restoration was classified as acceptable or non-acceptable (with the latter indicating that the restoration must be repaired or replaced).

Training for using FDI criteria was performed only after the first restorations' evaluation in an attempt to avoid the influence of the knowledge of the FDI criteria in the evaluation based on personal judgment. The students underwent a total of 8 hours of specific training involving theoretical explanations, discussions, and assessment of clinical cases that were representative of each score of the FDI criteria. A benchmark examiner (R.O.R.) was responsible for the training session and had been trained and calibrated for using the criteria. After these procedures, the students and benchmark examiner evaluated restorations in 20 photographs to evaluate the interexaminer reproducibility. The evaluations in the children included in the study began only when the interexaminer weighted kappa value reached values greater than 0.75.

Before the examinations, the teeth were carefully cleaned with a rotating bristle brush and pumice/water slurry. Visual inspection was performed with the subjects positioned in a dental unit with operating light illumination, using a 3-in-1 syringe, plane dental mirror, and World Health Organization periodontal probe (Hu-Friedy, Rio de Janeiro, Brazil). All restorations were assessed independently and randomly distributed in the two assessments to avoid memory bias. Students were instructed to

examine the restorations based on the clinical characteristics of the restorations and not by the lifespan of primary teeth. The duration of each examination was measured using a digital stopwatch for all evaluations.

### Reference standard

After the examinations, two examiners who had experience in the assessment of restorations' quality and in using the FDI criteria (R.O.R. and T.L.L.) also examined the children in a joint session and restorations were classified by consensus according to the need for intervention: no intervention, repair, and replacement. Examiners' evaluation results were compared to the reference standard, and the non-coincident results were considered as false results.

### Statistical analyses

Data analyses were performed using STATA 13.0 software (Stata Corp., College Station, Texas, USA). The unit of the analyses was the examiners' assessment. Initially, a descriptive analysis was performed. The chi-square test was used to test the impact of using FDI criteria on the treatment decision compared with personal judgment, considering three conditions regarding assessments when using FDI criteria in comparison with previous personal judgment: to have no change between both assessments, to choose a more invasive approach using FDI or to choose a more invasive approach when using FDI. Then, frequencies of choice for no change versus less (or more) invasive approaches were compared, and confidence intervals (CI) adjusted per clustering (student as cluster variable) were calculated. We also presented these distributions considering subgroups of undergraduate and graduate students. As multiple comparisons were performed using the same data, but outcomes were not statistically correlated (rho=0.4, p=0.06), Bonferroni correction was used and the p-value was set as 0.025 for compensation [12]. These assessments were categorized for further analyses considering two possible outcomes: to be (vs. not be) less invasive compared with personal judgment (primary outcome) or to be (vs. not be) more invasive (secondary outcome) based on the FDI criteria. Using both outcomes together we aimed to explore different changes direction that may occur when using the FDI criteria. For those cases in which changes had been observed when using the FDI criteria, we

calculated the frequencies of cases in which there was 1. an agreement with reference examiner, 2. probable false negatives (reference examiner recommended more intervention than general examiners) or 3. false positives (reference examiner recommended less intervention than general examiners). After that, multilevel multiple Poisson regression analyses were used to identify possible associated factors with the mentioned outcomes. The levels considered were the examiner (distal) and the assessment (proximal). As independent variables are tested variables related to both levels: examiner level - time spent with FDI criteria (minutes, continuous variable), examiners' experience (reference standard vs. undergraduate or graduate students, categorical variable), or assessment (restoration) level – child's decayed, missing, and filled teeth (*dmf-t*) index (discrete variable), dental arch (superior vs. inferior, categorical variable), number of restored surfaces (single or multi-surface) and reference clinical decision-making (none intervention, repair or replacement). A level of significance of 0.20 in the unadjusted analyses was regarded for variables entry into the model, and a level of 0.05 was considered to be retained in the final models. Prevalence ratios (PRs) were calculated with 95% confidence intervals (CIs).

## RESULTS

Six (54.5%) boys and five (45.5%) girls, with a mean age of 6 years (standard deviation (SD) = 1.1), participated in the study. According to the reference standard assessment, 7 (25.9%) restorations were clinically satisfactory, 14 (51.9%) were reparable, and 6 (22.2%) were required replacement.

The children presented a mean *dmf-t* index of 6.3 (SD=2.8). The mean times to evaluate all restorations based on the FDI criteria were 34.4 min (SD=7.1) and 31.6 min (SD=4.6) for undergraduate and graduate students, respectively. The mean times to evaluate all restorations based on personal judgment were 15.2 min (SD=4.3) and 7 min (SD=0.7) for undergraduate and graduate students, respectively.

Table I shows the descriptive analysis of the restorations' evaluation by undergraduate and graduate students. In most cases, there was no difference on decision-making when the

restorations were evaluated based on FDI criteria and personal judgment, irrespective of examiners' experience, both compared with less or more invasive changes. Furthermore, there was a high concordance with the evaluation of the reference standard. Graduate students achieved a higher level of agreement with the reference standard compared to the undergraduate ones (Table I).

When decision-making was more invasive with FDI criteria than personal judgment, evaluations coincident with reference standard were lower (Tables I and II). A similar finding was observed for those cases in which a less invasive approach was chosen by using the FDI criteria (Tables I and II).

Table II shows the distribution of the results according to the reference standard considering the two outcomes. Irrespective of examiners' experience, there was a trend of false results (compared to the reference examiners) when a change in the treatment decision was registered

by using the FDI criteria. Similar rates of false results (not coincident with reference examiners) were observed when, using FDI, students changed their assessment both for a more or less invasive approach.

Multilevel multiple Poisson regression analyses are summarized in Table III. Examiners who spent more time for evaluation based on the FDI criteria were more invasive (PR=1.001, 95%CI=1.0001-1.002; p=0.03). Moreover, examiners were more invasive using FDI criteria when examining children with higher $dmf\text{-}t$ (PR=1.16, 95%CI=1.01-1.32; p=0.03). Conversely, examiners chose a less invasive option when assessed multi-surface with FDI criteria (PR=2.04, 95%CI=1.03-4.05; p=0.04). Despite not significantly associated, we observed a high frequency of restorations requiring repair among those cases in which a less invasive approach was chosen using FDI criteria (15 out of 26 cases, 58%).

**Table I -** Descriptive analysis of the restorations' evaluation by undergraduate and graduate students. N (%, 95 confidence interval (CI) adjusted per clustering- cluster variable:examiner)

| Examiner | No difference between FDI criteria and personal judgment | % Agreement with reference | More invasive using the FDI criteria in relation to personal judgment | % Agreement with reference | p value* | Less invasive using the FDI criteria in relation to personal judgment | % Agreement with reference | p value* | Total |
|---|---|---|---|---|---|---|---|---|---|
| Reference standard | 21 (78%) | -- | 2 (7%) | | -- | 4 (18%) | -- | | 27 |
| Undergraduate students | 99 (73%;CI:61%-83%) | 73.7% (CI:71%-76%) | 18 (13%; CI:4%-35%) | 5.6% (CI:2%-15%) | -- | 18 (13%; CI:7%-25%) | 5.5% (CI:0.05%-39%) | | 135 |
| Graduate students | 102 (76%; CI:60%-87%) | 79.2% (CI:78%-80%) | 18 (11%; CI:9%-20%) | 5.6% (CI: 0.4%-43%) | | 15 (11%; CI:5%-25%) | 27% (CI:12%-49%) | | 135 |
| Total | 222 (75%; CI:66%-82%) | 76% (CI:74%-79%) | 38 (13%; CI:7%-22%) | 5.6% (CI: 1%-19%) | <0.001 | 37 (12%; CI:8%-20%) | 15% (CI:6%-32%) | <0.001 | 297 |

*p value: Chi-square test considering frequencies related to the impact of using the FDI. As multiple comparisons were performed, Bonferroni correction should be considered (i.e., statistically significant differences were considered when p<0.025). To subgroups exploratory analyses (undergraduate vs. graduate students), we considered the interpretation of CIs. Dark grey cells symbolizes confidence intervals that do not overlap considering the groups – probably real differences observed. Light grey cells symbolizes central figures that seems to be different (tend to) but the confidence intervals overlapped between groups and may represent a lack of statistical power to demonstrate the difference – possible trends that should be interpreted with caution. No shadows cells suggests similar values (similar central values and CIs)

**Table II -** Distribution of the results according to the reference standard based on outcomes (to be less or more invasive using FDI criteria). N (%, 95 confidence interval (CI) adjusted per clustering- cluster variable:student)

| | Less invasive using the FDI criteria in relation to personal judgment | | | | More invasive using the FDI criteria in relation to personal judgment | | | |
|---|---|---|---|---|---|---|---|---|
| | In agreement with reference | Probable false negative* | Probable false positive** | Total | In agreement with reference | Probable false negative* | Probable false positive** | Total |
| Undergraduate students | 2 (5.5%;CI: 0.05%-39%) | 16 (89%; CI:39%-99%) | 0 (0%) | 18 | 1 (5.6%; CI:2%-15%) | 0 (0%) | 17 (94%; CI:85%-98%) | 18 |
| Graduate students | 4 (27%; CI:12%-49%) | 10 (67%; CI:38%-89%) | 1 (6%; CI:0.06%-43%) | 15 | 1 (5.6%; CI:0.4%-43%) | 0 (0%) | 17 (94%; CI:57%-99%) | 18 |
| Total | 6 (15%; CI:6%-32%) | 26 (79%; CI:53%-92%) | 1 (3%; CI:0.02%-26%) | 33 | 2 (5.6%; CI:1%-19%) | 0 (0%) | 34 (94%; CI:81%-98%) | 36 |

*Better than scored by reference; **Worse than scored by reference.

**Table III -** Multilevel multiple Poisson regression analyses to identify possible associated factors with outcomes: to be more or less invasive based on the FDI criteria in relation to personal judgment

| Variables | PR$_{crude}$ (95%CI) | p-value | PR$_{adjusted}$ (95%CI) | p-value | Variables | PR$_{crude}$ (95%CI) | p-value | PR$_{adjusted}$ (95%CI) | p-value |
|---|---|---|---|---|---|---|---|---|---|
| More invasive with FDI criteria | | | | | Less invasive with FDI criteria | | | | |
| *Examiner* | | | | | *Examiner* | | | | |
| Reference standard | 1 | | | | Reference standard | 1 | | | |
| Undergraduate students | 0.61 (0.10-3.80) | 0.60 | | | Undergraduate students | 0.82 ( 0.34-2.00) | 0.67 | | |
| Graduate students | 1.10 (0.44-2.78) | 0.84 | | | Graduate students | 1.15 (0.27-4.93) | 0.85 | | |
| *dmf-t* | 1.16 (1.02-1.32) | 0.02 | 1.16 (1.01-1.32) | 0.03 | *dmf-t* | 1.05 (0.92-1.20) | 0.49 | | |
| *Type of arch* | | | | | *Type of arch* | | | | |
| Superior | 1 | | | | Superior | 1 | | | |
| Inferior | 0.65 (0.30-1.42) | 0.28 | | | Inferior | 0.75 (0.33-1.70) | 0.48 | | |
| *Number of restored surfaces* | 1.58 (1.02-2.44) | | | | *Number of restored surfaces* | | | | |
| One | 1 | | | | One | 1 | | 1 | |
| Two or more | 1.24 (0.65-2.35) | 0.52 | | | Two or more | 2.00 (1.05-3.82) | 0.04 | 2.04 (1.03-4.05) | 0.04 |
| *Time spent with FDI criteria (minutes)* | 1.001 (1.0001-1.002) | 0.03 | 1.001 (1.0001-1.002) | 0.03 | *Time spent with FDI criteria (minutes)* | 1.00 (1.00-1.0004) | 0.14 | 1.00 (1.00-1.0004) | 0.14 |
| *Decision-making* | | | | | *Decision-making* | | | | |
| None intervention | 1 | | | | None intervention | 1 | | | |
| Repair | 1.33 (0.61-2.91) | 0.47 | | | Repair | 1.57 (0.70-3.53) | 0.27 | | |
| Replacement | 1.42 (0.55-3.69) | 0.47 | | | Replacement | 1.40 (0.04-0.18) | 0.52 | | |

Abbreviation: *dmf-t* = decayed, missing, and filled teeth index; PR = prevalence ratio.

## DISCUSSION

In recent years, the use of FDI criteria in clinical trials has increased significantly, accounting for 50% of the published papers in 2016 [13]. The FDI criteria have been described as practical, relevant, and standardized criteria [13]. Through these criteria, the restorations are examined according to different parameters and then classified as acceptable or non-acceptable). Therefore, we hypothesized that their use could make the decision-making less intuitive, aiding in the assessment of restorations' quality. This study was the first study that investigated the impact of the FDI criteria for evaluating restorations on examiners' decision-making compared with their previous personal judgment in primary teeth.

Due to the relatively short lifespan of primary teeth, the decision for re-intervention in a defective restoration is not always an easy task. However, in our study, restorations placed in primary molars of children with a mean age of 6 years were evaluated, i.e., that should remain

functional for at least 2 years. Thus, examiners were instructed to choose the treatment based on the clinical characteristics of the restorations and not by the lifesfan of primary teeth.

Differently from previously expected, the impact of using the FDI criteria was relatively low, considering that most results were not changed when using the criteria. On the other hand, we observed a high agreement between students and the reference standard. At this point, it is essential to emphasize that a sample of students was selected (invited) to participate in this study. This probably influenced this agreement, since the students' interest is proportional to the dedication on assessments. All examiners were trained to use the FDI criteria after the first evaluation of the restorations based on personal judgment to avoid a possible effect residual from one method over the other. Although no previous experience in using FDI, these students may present a particular behavior compared to the average. Representativeness may be related to the entire group of eligible learners, for

example, an entire school class [14]. This aspect; however, does not invalidate the assessment and a convenience sample is usually necessary as the first step to test the implementation of an approach as a learning strategy. The reference method may also be contributing to this finding. Most published studies on decision-making related to restorations did not use a clinically relevant reference method [15]. Using the expert's consensus as a reference may result in less impact on the implementation of the criteria, since similar domains may be used. On the other side, it may approximate the results from those is really important to be seen in a clinical situation and considered in the decision-making.

Besides the idea of capturing if the use of FDI criteria may change the decision-making process when examining restorations, we intended to investigate how this could happen. That is why, in our study, two outcomes were considered: FDI criteria being less and more invasive than personal judgment. Similar rates of change when using the criteria both for choosing more or less invasive conducts. Besides, the impact when using FDI criteria was mainly associated with false results when the reference examiners' assessments were considered as the correct answers. It could represent they would have the same potential (around 15%) of promoting the overtreatment or neglecting the ideal treatment when using the criteria.

Examiners who spent more time on evaluation with FDI criteria were more invasive. Overall, clinicians are more accustomed to drawing on previous experiences to establish a diagnosis [16]. This process is less-time consuming and more practical to perform as part of a daily clinical routine [16]. Hypothetical-deductive models of clinical reasoning, such as the FDI criteria, involve a stepwise process and, therefore, are more time-consuming. Clinicians consider more information and different possibilities to arrive at a diagnosis and the respective treatment decision [16]. In some cases, using a new index may lead to seeking more actively for more alterations, also influencing time spent for examination and the increase of false results.

The mean time to evaluate the restorations using the FDI criteria was approximately 2 times greater than that spent with personal judgment, irrespective of the examiners' experience level. A more time-consuming process may lead to more

doubts and then the restorations' evaluation could be more prone to errors influenced by external factors such as patients' oral health conditions. On the other hand, the evaluation based on the personal judgment was less time-consuming. Furthermore, it was influenced by the examiners' experience because undergraduate students spent more time performing the evaluations than graduate students.

Students tended to intervene more in the restorations using the FDI criteria when examined children with higher *dmf-t*, i.e. examiners tended to overestimate the decision-making in children with higher caries experience due to their worse oral health conditions. This may be attributed to cognitive bias related to the individuals' mental processing [17], which likely occurs when clinicians filter the information available according to their own experiences and beliefs [18]. Previous studies reported that the children's caries experience influenced the performance of visual inspection in detecting carious lesions in primary teeth, evidencing the occurrence of cognitive biases [19,20]. These results may be extrapolated to a population of children seeking dental treatment and with previous caries experience (presence of restorations), what is a result of the selection process, in which the consecutive sample of included children may reflect this profile of children seeking dental treatment.

In contrast, examiners were approximately twice less invasive when multi-surface restorations were evaluated with FDI criteria in comparison with personal judgment. Several parameters proposed by the FDI, each expressed in five scores (three acceptable and two unacceptable), can be challenging to evaluate in proximal restorations due to arch position. Recurrent caries in the cervical wall may be difficult to be detected and may explain this examiners' more conservative assessment. This assumption is corroborated by a high percentage of assessments that tend to recommend fewer interventions (repair or replacement) compared to the reference standard. Indeed, even not statistically significant, restorations requiring repair tended to be underscored by a relevant part of the students.

Overall, the examiners' experience did not influence clinical decision-making. In our study, undergraduate students were in their last year of study in the course of Dentistry and graduate

students had at least 2 years of experience, which may explain the similar performances in the evaluation of the restorations' quality. Furthermore, a detailed description of five possible conditions of each parameter examined using the FDI criteria was provided for the examiners. Consequently, the evaluations would be less influenced by the examiners. Similar results were found in some studies that tested the influence of examiner experience on the performance of visual inspection for detection of carious lesions after a standard training [7,21,22].

On the other hand, graduate students showed a closer approximation to reference standard. These more experienced students seemed to be more inclined incorrectly choosing less invasive management when using the FDI criteria. The experience may be a differential when visual inspection for caries detection is the deal [23]. In our case, some additional years of clinical experience and practice may be underlying this observed trend. Certainly, we have some limitations in these inferences due to it to is a result found in a subgroup analysis and present an exploratory nature. In addition, some possible flaws may have been included, regardless planned a priori, as inflated false-positive rates, chance differences in observed treatment effects, and low power for the comparisons of interest. This reinforces that the results based on subgroup analyses should be interpretated cautiosly [24].

Clinical diagnosis is a subjective process, and therefore susceptible to different interpretations depending on whether examiners are more or less conservative. These students may present several differences considering a general sample of undergraduate and graduate students, including individual skills such as knowledge, interest, and practical abilities. As stated, a convenience sample of restorations and examiners, which may have influenced the findings. Indeed, our sample may not reflect the actual distribution of scores from restorations in children who seeking dental treatment [11] and they may include sample spectrum bias (patient selection domain) [15] if the idea is to test the accuracy of the diagnostic method. As in our study, the main purpose was not to test the accuracy of methods, but to investigate their impact on examiners' decision-making, this problem was not eliminated, but certainly was minimized. Our sample tended to be more challenging than a sample collected consecutively among patients seeking dental

treatment [11]. The idea during sample selection was intentionally to create situations in which we could better observe the actual impact of using FDI system (if there was), e.g., different examples of restorations demanding any type of intervention. On the other hand, despite using a convenience sample, we rejected our null hypothesis, since less or more invasive treatment when using FDI were significantly less frequent than no changes in decision-making process by using the criteria. Therefore, we can conclude we had statistical power to test our hypothesis. Therefore, some inferences were very definitive and may open discussion about some important points when using FDI criteria.

In general, the FDI criteria seemed to not always contribute to less invasive conduct among students, and some clinical experience may be required to reveal its actual benefit in assessing restorations (restored dental surfaces). Besides, its use among students increased the overtreatment, which could be related to a more actively seek for those proposed items in the criteria checklist. Therefore, the use of these criteria seems to cause more problems than bringing solutions to the studied group. Although the use of the FDI criteria seems to be logical, clinicians (mainly more experienced ones) think differently from each other in different clinical settings. Further studies addressing the performance of examiners more experienced in the restorations' evaluation based on FDI criteria, or even, including a wider variety of students, are necessary.

However, we identified specific points that could be useful when proposing the use of the criteria. Some recommendations may be included in the training process, and it could benefit the final implementation in a clinical setting. A non-so-active search for failures related to restorations and a conscious exercise of not being influenced by patients' caries experience may be useful tools to be recommended for those students who are interested in using the FDI criteria. This approach may minimize possible negative results that a non-systematic initiation could bring. These aspects may be the target of initiatives for educational purposes.

## CONCLUSION

The FDI criteria negatively influenced the restorations' evaluation and treatment decisions.

A more time-consuming process for evaluation and higher children's caries experience may lead to more invasive decision-making using the FDI criteria. In contrast, the evaluation of multi-surface restorations resulted in lower frequency of intervention based on the FDI criteria. The decision-making process was not affected by examiners' experience.

## Acknowledgements

## Authors' Contributions

DP: methodology, writing - original draft preparation. CPC: methodology, writing - review & editing. ROR: methodology, writing - review & editing. MMB: formal analysis. TMA: writing - review & editing. TLL: conceptualization, project administration, writing - review & editing.

## Conflict of Interest

The authors have no proprietary, financial, or other personal interest of any nature or kind in any product, service, and/or company that is presented in this article.

## Funding

## Regulatory Statement

This study was conducted in accordance with all the provisions of the local human subjects oversight committee guidelines and policies of the Federal University of Santa Maria. The approval code for this study is CAAE: 77219817.9.0000.5346.

## REFERENCES

1. Hilgert LA, Frencken JE, de Amorim RG, Mulder J, Leal SC. A study on the survival of primary molars with intact and with defective restorations. Int J Paediatr Dent. 2016;26(5):383-90. http://dx.doi.org/10.1111/ipd.12215. PMid:26567086.

2. Chisini LA, Collares K, Cademartori MG, Oliveira LJC, Conde MCM, Demarco FF, et al. Restorations in primary teeth: a systematic review on survival and reasons for failures. Int J Paediatr Dent. 2018;28(2):123-39. http://dx.doi.org/10.1111/ipd.12346. PMid:29322626.

3. Pedrotti D, Ribeiro JF, Pires CW, Rocha RO, Ardenghi TM, Soares FZM, et al. Survival and associated risk factors of resin-based composite restorations in primary teeth: a clinical, retrospective, university-based study. Pediatr Dent. 2017;39(4):313-8. PMid:29122073.

4. Wilson N, Lynch C, Brunton P, Hickel R, Meyer-Lueckel H, Gurgan S, et al. Criteria for the Replacement of Restorations: Academy of Operative Dentistry European Section. Oper Dent. 2016;41(S7):S48-57. http://dx.doi.org/10.2341/15-058-O. PMid:27689930.

5. Gordan VV, Riley J 3rd, Geraldeli S, Williams OD, Spoto JC 3rd, Gilbert GH. The decision to repair or replace a defective restoration is affected by who placed the original restoration: Findings from the National Dental PBRN. J Dent. 2014;42(12):1528-34. http://dx.doi.org/10.1016/j.jdent.2014.09.005. PMid:25223822.

6. Kanzow P, Wiegand A, Göstemeyer G, Schwendicke F. Understanding the management and teaching of dental restoration repair: systematic review and meta-analysis of surveys. J Dent. 2018;69:1-21. http://dx.doi.org/10.1016/j.jdent.2017.09.010. PMid:28943362.

7. Bussaneli DG, Boldieri T, Diniz MB, Rivera LML, Santos-Pinto L, Cordeiro RDCL. Influence of professional experience on detection and treatment decision of occlusal caries lesions in primary teeth. Int J Paediatr Dent. 2015;25(6):418-27. http://dx.doi.org/10.1111/ipd.12148. PMid:25511642.

8. Wilson N, Lynch CD, Brunton PA, Hickel R, Meyer-Lueckel H, Gurgan S, et al. Criteria for the replacement of restorations: Academy of Operative Dentistry European Section. Oper Dent. 2016;41(S7):S48-57. http://dx.doi.org/10.2341/15-058-O. PMid:27689930.

9. Hickel R, Peschke A, Tyas M, Mjör I, Bayne S, Peters M, et al. FDI World Dental Federation - clinical criteria for the evaluation of direct and indirect restorations. Update and clinical examples. J Adhes Dent. 2010;12(4):259-72. http://dx.doi.org/10.1007/s00784-010-0432-8. PMid:20847997.

10. Loguercio AD, Paula EA, Hass V, Luque-Martinez I, Reis A, Perdigão J. A new universal simplified adhesive: 36-month randomized double-blind clinical trial. J Dent. 2015;43(9):1083-92. http://dx.doi.org/10.1016/j.jdent.2015.07.005. PMid:26159382.

11. Moro BLP, Freitas RD, Pontes LRA, Pássaro AL, Lenzi TL, Tedesco TK, et al. Influence of different clinical criteria on the decision to replace restorations in primary teeth. J Dent. 2020;101:103421. http://dx.doi.org/10.1016/j.jdent.2020.103421. PMid:32615237.

12. Yoon FB, Fitzmaurice GM, Lipsitz SR, Horton NJ, Laird NM, Normand SLT. Alternative methods for testing treatment effects on the basis of multiple outcomes: simulation and case study. Stat Med. 2011;30(16):1917-32. http://dx.doi.org/10.1002/sim.4262. PMid:21538986.

13. Marquillier T, Doméjean S, Le Clerc J, Chemla F, Gritsch K, Maurin JC, et al. The use of FDI criteria in clinical trials on direct dental restorations: A scoping review. J Dent. 2018;68:1-9. http://dx.doi.org/10.1016/j.jdent.2017.10.007. PMid:29055692.

14. Cook DA, Reed DA. Appraising the quality of medical education research methods: the medical education research study quality instrument and the newcastle-ottawa scale-education. Acad Med. 2015;90(8):1067-76. http://dx.doi.org/10.1097/ACM.0000000000000786. PMid:26107881.

15. Signori C, Gimenez T, Mendes FM, Huysmans MCDNJM, Opdam NJM, Cenci MS. Clinical relevance of studies on the visual and radiographic methods for detecting secondary caries lesions

– A systematic review. J Dent. 2018;75:22-33. http://dx.doi.org/10.1016/j.jdent.2018.05.018. PMid:29857074.

16. Gowda D, Lamster IB. The diagnostic process. Dent Clin North Am. 2011;55(1):1-14. http://dx.doi.org/10.1016/j.cden.2010.08.002. PMid:21094715.

17. Evans JSBT. Dual-processing accounts of reasoning, judgment, and social cognition. Annu Rev Psychol. 2008;59(1):255-78. http://dx.doi.org/10.1146/annurev.psych.59.103006.093629. PMid:18154502.

18. Elstein AS. Thinking about diagnostic thinking: a 30-year perspective. Adv Health Sci Educ Theory Pract. 2009;14(Suppl. 1):7-18. http://dx.doi.org/10.1007/s10459-009-9184-0. PMid:19669916.

19. Moro BLP, Novaes TF, Pontes LRA, Gimenez T, Lara JS, Raggio DP, et al. The influence of cognitive bias on caries lesion detection in preschool children. Caries Res. 2018;52(5):420-8. http://dx.doi.org/10.1159/000485807. PMid:29566393.

20. Mendes FM, Novaes TF, Matos R, Bittar DG, Piovesan C, Gimenez T, et al. Radiographic and laser fluorescence methods have no benefits for detecting caries in primary teeth. Caries Res. 2012;46(6):536-43. http://dx.doi.org/10.1159/000341189. PMid:22907166.

21. Zandona AGF, Al-Shiha S, Eggertsson H, Eckert G. Student versus faculty performance using a new visual criteria for the detection of caries on occlusal surfaces: an in vitro examination with histological validation. Oper Dent. 2009;34(5):598-604. http://dx.doi.org/10.2341/08-082-L. PMid:19830976.

22. Gimenez T, Bittar D, Piovesan C, Guglielmi C, Fujimoto K, Matos R, et al. Influence of examiner experience on clinical performance of visual inspection in detecting and assessing the activity status of caries lesions. Oper Dent. 2013;38(6):583-90. http://dx.doi.org/10.2341/12-067-C. PMid:23617691.

23. Gimenez T, Piovesan C, Braga MM, Raggio DP, Deery C, Ricketts DN, et al. Visual inspection for caries detection: a systematic review and meta-analysis. J Dent Res. 2015;94(7):895-904. http://dx.doi.org/10.1177/0022034515586763. PMid:25994176.

24. Wang X, Piantadosi S, Le-Rademacher J, Mandrekar SJ. Statistical considerations for subgroup analyses. J Thorac Oncol. 2021;16(3):375-80. http://dx.doi.org/10.1016/j.jtho.2020.12.008. PMid:33373692.

**Tathiane Larissa Lenzi**
**(Corresponding address)**
Universidade Federal do Rio Grande do Sul, School of Dentistry, Post-Graduate Program in Pediatric Dentistry, Porto Alegre, RS, Brazil.
Email: tathilenzi@hotmail.com