Turbidity tubes for drinking water quality assessments
Caetano C. Dorea and Murray R. Simpson

ABSTRACT

Turbidity tubes have been considered to be the field method of choice for drinking water quality monitoring in resource-limited contexts because of their relative simplicity and low cost in comparison with conventional (nephelometric) turbidimeters. These tubes utilise the principle of visual extinction of a submerged target for turbidity determination and were therefore thought to be subject to user subjectivity, possibly affecting results. This study evaluated their performance under both field and controlled-laboratory conditions. Results from turbidity tubes can differ substantially from those obtained with conventional turbidimeters; this is of particular importance in the reporting of low turbidity (<10 NTU) measurements. These differences could be due to a combination of factors, such as: user variability, differences in calibration scales, and turbidity tube target shape and background colour. In view of their limitations, the usefulness of turbidity tubes for drinking water quality assessments and recommendations on the reporting of their results are also discussed.

Key words | aesthetics, transparency, tube, turbidity, visual extinction

INTRODUCTION

In developing countries, water quality assessments of critical parameters such as turbidity play an essential role in determining the necessary level of treatment and attainment of drinking water quality standards. Turbidity can be defined as an ‘expression of the optical property that causes light to be scattered and absorbed rather than transmitted with no change in direction or flux level’ through a sample (Standard Methods APHA et al. 2005). Such light scattering is typically caused by waterborne particles that can be of mineral or organic origin (i.e. suspended solids).

Although there are no direct health implications associated with turbidity in water, particles causing ‘high levels of turbidity’ are considered to possibly interfere with disinfection (WHO 2008a). The proposed mechanisms for interference vary and depend on the physical and chemical characteristics of the turbidity-causing suspended particulates, thus making generalisations difficult. Furthermore, turbidity can also influence user acceptability and is therefore considered to be an important aesthetic parameter. This is an important consideration, particularly in developing country contexts. For this purpose, WHO (2008a) has suggested that waters with less than 5 nephelometric turbidity units (NTU) are usually considered acceptable by consumers. However, a median turbidity of less than 0.1 NTU is recommended to ensure effective disinfection.

Modern turbidity measurements are usually made with nephelometers, in which turbidity is generally correlated to the intensity of light scattered at a 90° angle from the incident light (Sadar 1998). These turbidimeters can be relatively expensive (approximately US$2,000) and typically require calibration with standards (e.g. formazin, latex spheres) with associated costs and relevant laboratory infrastructure, thus further proscribing their use in many resource-limited contexts. In such circumstances, turbidity tubes have been used for field measurements because of their relative simplicity and relative low cost (approximately US$80). Here, turbidity determinations are based on the principle of visual extinction, whereby the sample is intermittently poured into a transparent tube (typically 600 to 650 mm long) with a pre-determined scale on its side and a distinct pattern (e.g. crosshair, circle or alternate black and white ‘Secchi’ quadrants) etched on its bottom. At each pouring
interval the reader looks down the tube (Figure 1) and continues to pour the sample until the pattern can no longer be seen; that is, it becomes ‘extinct.’ The water level inside the tube is then measured against the scale on its side to estimate the turbidity.

Field experience has shown that there can be (significant) discrepancies between measurements on the turbidity tubes and determinations performed on electronic turbidimeters. The objectives of this paper are to provide background to the development of these useful field tools as well as to provide a first assessment of their performance and identification of their limitations for drinking water quality applications.

BACKGROUND

Turbidity tubes have been derived from the early diaphanometers (from the Latin diaphanus, or Greek diaphanous, meaning transparent) used to measure turbidity, which was used as a surrogate for the rapid estimation of suspended matter. These instruments consisted of a metal tube with a partially painted glass bottom on which the unpainted portion formed a cross of two lines (Parmelee & Ellms 1899). Light was reflected up through the tube and water was poured in until the cross image on the bottom disappeared from sight. The depth of the water column in the tube was noted and the turbidity calculated as the reciprocal of the depth. Typically, these determinations were then converted into estimates of (gravimetric) suspended matter concentrations through pre-determined regressions.

A subsequent development was the Jackson candle turbidimeter (Figure 2), where, in an attempt to standardise the light source, a candle was used. It was one of the most popular visual extinction methods and was similar in principle to the diaphanometer, the difference being that the glow of a candle was the visual target end-point and readings were reported in Jackson turbidity units (JTU). A further refinement was the Parr turbidimeter, in which an electric light source replaced the candle (Vanous et al. 1982). Although the Jackson candle turbidimeter and diaphanometers were thought to work well and provide reasonably accurate readings (at the time), improvements in early turbidity measurements came from the development of standard turbidity solutions using finely divided silica from diatomaceous earth to calibrate instruments (Whipple & Jackson 1933). The ppm-silica scale prevailed until the middle of the 20th century, after which formazin was adopted as the standard because of its advantages in terms of reproducibility, light-scattering characteristics and traceability (Sadar 1998). Details of the wire method (Figure 2) and other early methods of turbidity measurement of historical interest can be found in the aforementioned references.

Existing turbidity tubes used for drinking water quality monitoring are based on the design originally developed for the DelAgua water quality test kit in the early 1980s (Lloyd et al. 1985). At the time, shorter tubes (300 mm long) were commercially available, but were only marked for values above 20 turbidity units (TU). This was clearly not suitable for use alongside the maximum WHO drinking water turbidity guideline value of 5 NTU and as a result the longer tube (650 mm) format was crafted (B.J. Lloyd, personal communication, 2011). The turbidity scales on such tubes were determined through calibrations based on visual extinction measurements done on 13 pre-prepared commercially available turbidity standards ranging from 5 to 2,000 NTU. The mean of the water column heights
determined by volunteers for each standard were used as the template to develop the calibration scale to be engraved on the turbidity tubes.

Transparency or clarity tubes are a contemporary variation of turbidity tubes used mainly for surface water quality assessments. Although there are different commercially available models of such tubes that vary in size and format, they are based on the same visual extinction principles as the turbidity tubes. One particular difference is that in many transparency tubes the ‘transparency’ or ‘clarity’ readings are taken as the water column depth and subsequently correlated with total suspended solids and turbidity. An interesting variation encountered in transparency tubes is a valve placed at the bottom of the tube that allows water to be drained until the target is just barely visible. Characterisation tests of transparency tubes for in vitro turbidity assessments have yielded satisfactory results demonstrating their suitability for their required purposes (Anderson & Davic 2004; Dahlgren et al. 2004). However, such assays are primarily intended for monitoring of surface waters in which quality criteria are relatively higher and more variable (e.g. 10–150 NTU; LDEQ 2007; NRB 2008) than drinking water guideline values.

METHODS

Turbidity tube measurements were made on four commercially available turbidity tubes. Their characteristics are summarised in Table 1. The calibration scale of the DelAgua was determined as described previously. The manufacturers of the Prerana model indicated that the same (i.e. DelAgua) scale was adopted on their tubes. There was no indication by the other manufacturers or in published literature as to how the other turbidity scales were determined. A comparison of the tubes (Figure 3) suggests that such scales, albeit with different ranges, were determined in a similar fashion or were reproduced (i.e. Prerana) from an original scale with the exception of one specimen (i.e. Cole-Parmer). As each tube had a different target at its bottom, such
similarities in scale allowed for the evaluation of the target type on visual extinction determination and user preference.

Although the studied turbidity tubes had scales calibrated as NTU, reporting turbidity tube results as such could be misleading as the turbidity is not being estimated by true nephelometry, rather through visual extinction of a target (i.e. light attenuation). Nephelometric (‘actual’) turbidity measurements were reported in NTUs and were performed with a 2100P turbidimeter (HACH, USA) that was calibrated according to the manufacturer’s instructions. Turbidity tube readings were reported in diaphanometric turbidity units (DTUs). This proposition is thought to better reflect the method used and it also differs from JTUs, which are determined using an artificial light source (i.e. Jackson candle).

Turbidity tube measurements in the field and laboratory were performed by slowly pouring the water sample along the side of the turbidity tube to avoid the formation of air bubbles, which could affect the visualisation of the target. The intermittent filling of the tube was done until the target could no longer be seen, with the observer holding the tube vertically by its bottom portion. The depth of the water column in the tube was compared with the tube’s (non-linear) turbidity scale, from which the turbidity was estimated. Tubes were first rinsed with the ‘clearest’ water available (e.g. 0.45 μm membrane-filtered rainwater, bottled water or deionised water) then with the water sample before taking readings. Care was taken not to stand facing direct sunlight and readings were taken in the observer’s shadow in field measurements. In laboratory tests, the visual extinction depth determined by the observer was also recorded with a tape measure and subsequently converted to DTUs using calibration scale plots (Figure 3). No distinction was made between observers with glasses or contact lenses.

In field measurements only the DelAgua tube was used. Samples were analysed from varied water sources in three different locations, namely: Scotland (pond and river water, with kaolin-augmented turbidities), Indonesia (water from springs and rainwater harvesting tanks), and Rwanda (treated river water sampled at different stages of treatment plant). Readings were taken by different observers (one for each location) and in varied light conditions (not recorded).

A laboratory-based volunteer experiment was conducted to compare the four turbidity tube models and to assess any differences between individual observers under controlled conditions. In their first use of turbidity tubes, 60 volunteers were given three different suspensions and determined their turbidities using each of the four turbidity tubes. Volunteers recorded the depth of the water column causing target extinction using a tape measure in order to avoid possible judgement errors with regard to the estimation of turbidities on the tubes’ (non-linear) scales. The three suspensions were made by diluting a formazin stock solution in deionised water, prepared in accordance with Standard Methods.

### Table 1 | Turbidity tube characteristics

<table>
<thead>
<tr>
<th>Tube manufacturer</th>
<th>Range (DTU)</th>
<th>Dimensions (mm)</th>
<th>Calibration scale best fit (ED: extinction depth in mm)</th>
<th>Target description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DelAgua (UK)</td>
<td>5 - 2,000</td>
<td>635 OD 30 ID 24</td>
<td>DTU = 74,664(ED)^{-1.483} (R^2 = 0.9997)</td>
<td>Empty black circle on a yellow base</td>
</tr>
<tr>
<td>Wagtech (UK)</td>
<td>5 - 500</td>
<td>648 OD 32 ID 26</td>
<td>DTU = 75,639(ED)^{-1.483} (R^2 = 0.9997)</td>
<td>Black crosshair on a yellow base</td>
</tr>
<tr>
<td>Prerana (India)</td>
<td>5 - 2,000</td>
<td>628 OD 32 ID 25</td>
<td>DTU = 53,362(ED)^{-1.423} (R^2 = 0.9984)</td>
<td>Empty black circle on a yellow base</td>
</tr>
<tr>
<td>Cole-Parmer (USA)</td>
<td>5 - 500</td>
<td>641 OD 32 ID 26</td>
<td>DTU = 310.02e^{-0.007(ED)} (R^2 = 0.9568)</td>
<td>Black crosshair on a transparent base</td>
</tr>
</tbody>
</table>

![Figure 3 | Turbidity tube calibration scale comparison.](https://iwaponline.com/washdev/article-pdf/1/4/233/384481/233.pdf)
(APHA et al. 2005), to low, medium and high nominal target turbidities of 6, 60 and 600 NTU, respectively. A pilot test was conducted with dilutions of a kaolin clay suspension, but finally the formazin solutions were opted for as they had a greater repeatability in addition to other advantages (Sadar 1998). Also, the use of formazin suspensions would be consistent with how at least one of the tubes was calibrated (i.e. DeAgua tube). Volunteer data was analysed through analysis of variation (ANOVA). Finally, volunteers were asked to indicate their preferred tube with regard to ease of target extinction determination.

RESULTS AND DISCUSSION

Field studies

Results from field measurements are presented in Figure 4. Owing to the turbidity tube’s lower limit on the calibration scale (5 DTU), as expected, turbidities of less than 5 NTU could not be determined on the tubes. The majority of low turbidity (i.e. <10 NTU) samples from the Scottish (river and pond water) and Indonesian (spring and stored rainwater) sites were too low for turbidity tube readings. However, field data from Rwanda suggests that such tubes can also under estimate actual turbidities by up to three times less than the actual NTU values, particularly in the lower measurements. In general, turbidity tube readings either over- or under estimated actual turbidity values for particular locations. These trends could be due to site-specific factors such as: suspension characteristics, lighting conditions and different observers.

The ‘higher’ turbidity Scottish samples were generated with the addition of kaolin. In contrast, samples from the Rwandan site consisted of river water undergoing treatment in a rural multi-stage filtration plant (Dorea et al. 2004). As treatment progressed through subsequent sedimentation and filtration stages, the turbidity-causing particles that remained in suspension were likely to be of a finer and finer nature. Early visual extinction work indicated that finer suspensions yielded smaller extinction depths (i.e. analogous to higher turbidity tube readings) than coarser suspensions for the same concentration of suspended matter (Parmelee & Ellms 1899), which is in agreement with current understanding of light-scattering of particle-laden suspensions (Sadar 1998). In addition, particle origin (i.e. mineral or biological) and dissolved substances (i.e. colour) could have influenced results, as has been observed to occur in transparency tube studies (Kilroy & Biggs 2002). However, it was not possible to further evaluate the trends in the Scottish and Rwandan samples with regard to suspension properties because of limited field techniques available at the time. Although lighting conditions were not recorded, they likely differed between sites. Other studies on visual transparency and clarity determination methods (Kilroy & Biggs 2002; Steele & Neuhausser 2002) showed no significant differences in readings performed in varied natural ambient lighting conditions (i.e. sunny vs. cloudy). It could, therefore, be plausible that the observed trends were also not affected by different lighting conditions. No inferences can be drawn on the influence of different observers on the obtained results. However, it is thought that the visual extinction assessment as well as the estimation of the turbidity value, typically to the nearest 10 DTU, on the (non-linear) tube calibration scale could be subject to user bias and may have affected the results.

Monitoring of treatment process performance is a useful application of turbidity tubes. However, data from the Rwandan site demonstrated that turbidity reduction efficiencies of different treatment stages, calculated using turbidity tube data, are over-estimated in comparison with those calculated using a portable nephelometric turbidimeter (Figure 5). This treatment efficiency over-estimation could be higher in the case of slow sand filter samples, as
efficiency calculations were made with the turbidity tube’s lowest calibration point (5 DTU) even though actual readings could have been less than 5 DTU.

Volunteer tests

The recorded extinction depths for each tube at the different target turbidity suspensions from the volunteer tests are summarised in Figure 6 on a logarithmic scale for clarity. An ANOVA indicated that there was no statistically significant uniformity in recorded visual extinction depths ($\alpha = 0.01$) between the different turbidity tubes. Further $t$-test analyses ($\alpha = 0.01$) revealed that certain turbidity tube pairs performed similarly at particular turbidity levels, but no significant correlation between this performance and differentiating tube characteristics (i.e. target pattern and base colour) were found. It should be noted, however, that many extinction depths in the low turbidity samples (i.e. 6 NTU) were limited by the turbidity tubes’ maximum depth (Table 1) and do not necessarily reflect the true observed extinction depth. The magnitude of the observed ranges decreased with increasing turbidities, which is in agreement with the findings of Anderson & Davic (2004) on transparency tubes of similar internal diameter. The volunteers were first-time turbidity tube users and it is possible that with experience such range in results could decrease.

Figure 7 summarises the results of an a posteriori conversion of the observed extinction depths (Figure 6) to the nearest turbidity unit using the tubes’ calibration scales (Figure 3), noting that values outside the marked calibration scales were extrapolated. At higher turbidities (i.e. 60 and 600 NTU), the calibration curves of the turbidity tubes resulted in under-estimates of the target turbidity, except for the Cole-Parmer tube tube in the 6 and 60 NTU determinations. This is thought to be due to its different calibration scale, which would over estimate turbidities in that range given a similar extinction depth (Figure 3). The results from 6 NTU formazin solution readings appeared to be more precise. However, the extinction depths were limited to the lengths of the tubes resulting in many readings of 5 DTU. This was the minimum possible reading in most tubes, except the Cole-Parmer tube which was 4 DTU (through extrapolation of the calibration scale). A different trend may have resulted with longer tubes (i.e. some
transparency tubes are up to 1,200 mm long). The range of turbidity tube measurements increased with higher turbidity values. As seen from the calibration scales (Figure 3), small variations in extinction depths can result in large variations in turbidity tube readings. Under the experimental conditions tested, an ANOVA analysis revealed that the tubes were able to differentiate with statistical confidence between turbidities of different orders of magnitude. Yet, at each tested turbidity, there was no statistically significant consistency of the turbidity determinations between the tubes. These differences in turbidity estimates were possibly due to the combined effect of errors from user variability (Anderson & Davic 2004), differences in calibration scales, and turbidity tube target shape and background colour.

Extinction depths could provide more reliable measurements, as it would eliminate the need to estimate turbidity value from the calibration scale. Such estimates could be a significant source of variation owing to the non-linear turbidity scale. However, for water quality and treatment applications, the extinction depths would still need to be converted to turbidity units for a more meaningful interpretation. Current calibration scales may not be well-suited for the variety of conditions faced in practice, as judged by field data (Figure 4). Further research on factors affecting visual extinction of targets could improve calibration scales.

As the tested tubes were similar in dimensions, volunteer preferences were interpreted in terms of the type of target shape and background. The Wagtech tube (with the black cross set on a yellow background) was rated similarly to the tubes with black circles set on a yellow background (DelAgua and Prerana), with 42% and 48% of preferences, respectively. However, only the remaining 10% chose the target consisting of a black cross etched on a transparent background (i.e. Cole-Parmer) as the easiest target on which to determine visual extinction. Anderson & Davic (2004) tested different targets (‘Secchi’ pattern and black dot on white background), but no clear volunteer preferences were identified. According to Davies-Colley (1988), theoretically, black targets should have an advantage over Secchi-type patterns for water clarity measurements. Carlson (2006) compared two types of transparency tube that differed in terms of their target patterns (Secchi and matt black) and tube reading direction (i.e. vertically versus horizontally held). He found that the two instruments were measuring the same aspects of water clarity. However, no conclusion could be drawn with regard to a certain target having a particular advantage over another. Moreover, targets commonly used in transparency tubes are different from those used in turbidity tubes, and their application to drinking water quality monitoring purposes requires further investigation.

**GENERAL DISCUSSION**

Turbidity tube measurements in field conditions and in controlled environments can vary significantly. Such variation can occur between different users analysing the same sample as well as between different locations on samples of the same nephelometric turbidity. The variability in turbidity tube readings is of particular significance in low-turbidity measurements (i.e. <10 NTU). Turbidity tubes were designed for measurements down to 5 NTUs in conformity with aesthetically based WHO guideline values, as turbidity is said to be initially noticed by the naked eye above approximately 4 NTU (WHO 2011). Nonetheless, turbidity tubes are frequently used as ‘pass/fail’ gauges of water quality in monitoring campaigns with regard to their 5 DTU limit (assumed to be equivalent to 5 NTUs). It has been shown that turbidity tube readings can result in under-estimates of actual values (Figure 4), possibly resulting in ‘false passes’ in monitoring exercises. Further research is warranted on the influence of suspension properties (e.g. particle origin, dissolved substances) on turbidity tube measurements.

As a result of the differences in conventional nephelometry and turbidity readings made with turbidity tubes, it is important to differentiate between the two methods when reporting results. For this reason, different units for turbidity tube measurements have been proposed. JTUs have sometimes been used for this purpose (WHO 2008b), but may not be an accurate portrayal as no fixed light source (i.e. Jackson candle) is used in turbidity tube measurements (typically ambient light). WHO (2008b) has stated that the two units (i.e. JTU for turbidity tubes and NTU for nephelometry) can be used interchangeably. If the calibration scales on turbidity tubes gave consistent
results with those obtained with conventional turbidimeters in varied conditions, interchangeable units could be acceptable. However, it has been shown that there can be substantial differences between the two methods of turbidity measurement. As such, the two units should not be used interchangeably. In fact, good practice in turbidity measurements dictates that, in addition to the utilisation of the relevant units, the instrument type and model should be specified as performance can vary (Gippel 1995); the same should apply to turbidity tube measurements. Arguably, the use of extinction depths could have advantages from the standpoint of avoiding errors due to turbidity scale ‘guestimates’. However, as these tubes are used for drinking water quality monitoring and process control, relating the readings to a known parameter (i.e. turbidity) is advantageous in terms of the interpretation of results by operators.

There is need for an accessible, simple and affordable method for routine turbidity measurements (Sobsey et al. 2009; WHO 2011). Lloyd & Helmer (1991) stated that it is more important to examine a water supply frequently with simple and reliable tests than only sporadically with more complex tests. To this end, turbidity tubes, despite their limitations, can still fulfil an important aspect of routine monitoring of water supplies in developing countries until a more adequate method is developed. Water safety plans (WSPs) have been proposed as an approach to improve water quality management in developing countries (Howard 2005). WSPs rely on process monitoring and control measures in all stages of production and distribution, from catchment to tap, to ensure water safety (Bartram et al. 2009). In the case of systems with surface water treatment plants utilising chemical coagulation and filtration, it is likely that turbidity will be one of the main operational parameters to be monitored. However, given the turbidity tube’s lowest stated measureable turbidity of 5 DTU and the variability of their readings, it is questionable whether such instruments can yield sufficiently reliable information for the implementation of WSPs in developing countries. In order to give more informative results for this type of monitoring, it would be desirable to be able to measure turbidities less than 5 DTU, requiring longer turbidity tubes (>600 mm). However, according to Kilroy & Biggs (2002), in vitro visual assays such as these can be unreliable in tube lengths greater than 600 mm.

Despite their limitations, turbidity tubes can still be very useful and low cost tools to aid in drinking water quality assessments in resource-poor contexts, yielding valuable information. For example, Lantagne (2008) recommended different sodium hypochlorite dosages for household and emergency disinfection of unimproved water sources depending on their turbidities, which were classified in three bands of low (less than 10 NTU), medium (between 10 and 100 NTU) and high (greater than 100 NTU). Based on volunteer tests, it is expected that turbidity tubes could differentiate between turbidities in such ranges and could therefore be useful in the design of such chlorination campaigns.

CONCLUSIONS

Turbidity tubes can yield results with significant differences when compared with turbidity measurements using electronic nephelometers. As such, the use of the unit DTU to identify readings made with turbidity tubes and differentiate them from other methods has been proposed. Significant variability between turbidity tubes measuring the same sample was also observed. They can, however, make a reasonable distinction between samples of different orders of magnitude.

ACKNOWLEDGEMENTS

Ashley Taylor and Lisa Vanhaute are acknowledged for their data collection efforts during preliminary work of this study. The British Council (for a summer intern through the IAESTE programme), the John Robertson Bequest (for equipment) and the Royal Geographical Society (for fieldwork) are thanked for their funding contributions to different aspects of this study.

REFERENCES


First received 17 August 2011; accepted in revised form 25 October 2011